Metabolism of steppe rivers in Mongolia and the United States: drivers, heterogeneity, methodological bias, and climate warming

by

Anne Eileen Schechner

B.A., Colby College, 2015

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Division of Biology
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2021

Abstract

River metabolism is a central component of global biogeochemical processes and has become a widely used metric of ecosystem function. With increasing data availability, broad-scale models of metabolism are now available towards prediction and improved understanding. Many river metabolism studies do not provide sufficient methodological detail for replication, but extrapolate from numerous local measurements and predict responses to various drivers, including climate. I was therefore interested in 1) how we can make metabolism estimates more accurate, representative, and comparable in methods and reporting? 2) what variables best explained temperate steppe metabolism and how did this vary by region and scale? and 3) how can we expect these rates to change under a warming climate? I use the answers to these questions to improve our understanding, reporting, and representativeness of studies of river metabolism.

In evaluating the reporting of open channel river metabolism methods, only 79% of 43 sampled papers published from 2015-2019 mentioned calibration, 44% described sensor placement, and 34% did not describe estimation approaches sufficient for replication. Given that spatial heterogeneity in rivers influences metabolism, and measurement sensitivities vary with sensor model, it is important to have appropriately established replicable protocols and detailed information in reported methods along with a holistic understanding of how river heterogeneity might influence metabolism. I deployed 2-8 sensors at 92 steppe river reaches to characterize site heterogeneity, evaluating how sensor placement and type, deployment length, drift correction, data source, local versus remotely sensed data, and calibration affected metabolism estimates. Estimates of gross primary production (GPP) and ecosystem respiration (ER) were strongly influenced by deployment location within a river reach; GPP and ER rates varied up to 131% and

69% respectively across a river width and up to two orders of magnitude within reach. Dissolved oxygen sensor brands vary widely in precision and accuracy; I found even when operated within stated performance ranges, estimates of GPP and ER could vary by 82% and 198% respectively if not properly calibrated, as determined using field data from a sample site. Inaccuracies from sensor drift over weeklong deployments led to 48% ER overestimation and 2% GPP overestimation comparing uncorrected with corrected field data. With a firmer understanding of methodological and riverine heterogeneity, I could more confidently compare our sites.

To explore explanatory structures across scales, I then linked metabolism estimates with reach-to-watershed-scale hydrogeomorphology, vegetation, climate, and anthropogenic impact metrics to evaluate predictors and applicability of traditional ecological frameworks in the Anthropocene. I expected that vegetation and climate related to ecoregion would be more explanatory than human or hydrogeomorphic data. I present the structures with the greatest explanatory power by river type, scale, and location. This required a systematic approach to identify the most explanatory variables, many of which were strongly correlated. I was subsequently interested in using these explanatory mixed models to predict change.

Responses of metabolism rates to climate change is critically important to global carbon cycling, so I used the above models to predict changes in GPP and ER under warmer temperatures. I evaluated the downscaling of broad-scale metabolism models using data collected from broad regions of Mongolia and North America. The understudied rivers of the semi-arid steppe of Mongolia are particularly vulnerable to climate change due to high altitude and latitude. This steppe has matching ecoregions with the United States Great Plains, allowing cross-continent investigation of temperature effects on river metabolism. I evaluate how a broad-scale modeling approach applied at the ecoregion level, projecting changes in estimated rates of

metabolism under different warming scenarios. Temperature was not the primary explanatory variable, but directly and indirectly influenced modeled rates of metabolism. Our metabolism models did not scale down well. The Grassland Steppe was the most temperature-sensitive ecoregion for both rates on both continents.

I offer best practices for more comparable, replicable, representative, and accurate methods in stream metabolism study, and present most explanatory structures of variables by river type, scale, and location. I conclude that macrosystem-scale studies require broader interdisciplinary and multi-scale assessment for prediction and capture of variation in aquatic metabolism, and the observed distribution of spatial patterns of river metabolism is scale-dependent. This suggests that researchers, managers, and policymakers must incorporate factors operating at multiple scales to inform management and climate projections, particularly if interested in modeling the influence of increased temperature on river metabolism.

Metabolism of steppe rivers in Mongolia and the United States: drivers, heterogeneity, methodological bias, and climate warming

by

Anne Eileen Schechner

B.A., Colby College, 2015

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Division of Biology
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2021

Approved by:
Major Professor
Walter K. Dodds

Copyright

© Anne Eileen Schechner 2021.

Abstract

River metabolism is a central component of global biogeochemical processes and has become a widely used metric of ecosystem function. With increasing data availability, broad-scale models of metabolism are now available towards prediction and improved understanding. Many river metabolism studies do not provide sufficient methodological detail for replication, but extrapolate from numerous local measurements and predict responses to various drivers, including climate. I was therefore interested in 1) how we can make metabolism estimates more accurate, representative, and comparable in methods and reporting? 2) what variables best explained temperate steppe metabolism and how did this vary by region and scale? and 3) how can we expect these rates to change under a warming climate? I use the answers to these questions to improve our understanding, reporting, and representativeness of studies of river metabolism.

In evaluating the reporting of open channel river metabolism methods, only 79% of 43 sampled papers published from 2015-2019 mentioned calibration, 44% described sensor placement, and 34% did not describe estimation approaches sufficient for replication. Given that spatial heterogeneity in rivers influences metabolism, and measurement sensitivities vary with sensor model, it is important to have appropriately established replicable protocols and detailed information in reported methods along with a holistic understanding of how river heterogeneity might influence metabolism. I deployed 2-8 sensors at 92 steppe river reaches to characterize site heterogeneity, evaluating how sensor placement and type, deployment length, drift correction, data source, local versus remotely sensed data, and calibration affected metabolism estimates. Estimates of gross primary production (GPP) and ecosystem respiration (ER) were strongly influenced by deployment location within a river reach; GPP and ER rates varied up to 131% and

69% respectively across a river width and up to two orders of magnitude within reach. Dissolved oxygen sensor brands vary widely in precision and accuracy; I found even when operated within stated performance ranges, estimates of GPP and ER could vary by 82% and 198% respectively if not properly calibrated, as determined using field data from a sample site. Inaccuracies from sensor drift over weeklong deployments led to 48% ER overestimation and 2% GPP overestimation comparing uncorrected with corrected field data. With a firmer understanding of methodological and riverine heterogeneity, I could more confidently compare our sites.

To explore explanatory structures across scales, I then linked metabolism estimates with reach-to-watershed-scale hydrogeomorphology, vegetation, climate, and anthropogenic impact metrics to evaluate predictors and applicability of traditional ecological frameworks in the Anthropocene. I expected that vegetation and climate related to ecoregion would be more explanatory than human or hydrogeomorphic data. I present the structures with the greatest explanatory power by river type, scale, and location. This required a systematic approach to identify the most explanatory variables, many of which were strongly correlated. I was subsequently interested in using these explanatory mixed models to predict change.

Responses of metabolism rates to climate change is critically important to global carbon cycling, so I used the above models to predict changes in GPP and ER under warmer temperatures. I evaluated the downscaling of broad-scale metabolism models using data collected from broad regions of Mongolia and North America. The understudied rivers of the semi-arid steppe of Mongolia are particularly vulnerable to climate change due to high altitude and latitude. This steppe has matching ecoregions with the United States Great Plains, allowing cross-continent investigation of temperature effects on river metabolism. I evaluate how a broad-scale modeling approach applied at the ecoregion level, projecting changes in estimated rates of

metabolism under different warming scenarios. Temperature was not the primary explanatory variable, but directly and indirectly influenced modeled rates of metabolism. Our metabolism models did not scale down well. The Grassland Steppe was the most temperature-sensitive ecoregion for both rates on both continents.

I offer best practices for more comparable, replicable, representative, and accurate methods in stream metabolism study, and present most explanatory structures of variables by river type, scale, and location. I conclude that macrosystem-scale studies require broader interdisciplinary and multi-scale assessment for prediction and capture of variation in aquatic metabolism, and the observed distribution of spatial patterns of river metabolism is scale-dependent. This suggests that researchers, managers, and policymakers must incorporate factors operating at multiple scales to inform management and climate projections, particularly if interested in modeling the influence of increased temperature on river metabolism.

Table of Contents

List of Figures

List of Tables

List of Equations

Acknowledgements

I received a tremendous amount of assistance and support throughout graduate school and the writing of this dissertation, and will attempt to thank those responsible here. I could not have asked for a better advisor than Dr. Walter Dodds, who helped me learn and grow as a limnologist and as a person. I don't think many people are lucky enough to have an advisor who gives so much help, patience, and guidance. And editing. I thank my advisory committee - Drs. John Blair, Lydia Zeglin, and Scott Kenner, as well as my outside chair, Eric Fitzsimmons - for supporting and shaping my work. I am extremely grateful for two unofficial mentors and advisors in the field, Flavia Tromboni and Alain Maasri, for their guidance, coauthorship, and collaboration. The MACRO project allowed me to collaborate with scientists around the world and a wealth of scientific disciplines, and brought me see more of this country than I ever had before graduate school. I am additionally grateful for the time I was able to spend in Mongolia, for the Mongolian researchers I learned from, and for all the Mongolians I was fortunate enough to meet. I have never known such hospitality and am very thankful for my time in the Land of the Blue Sky. баярлалаа.

I am indebted to my colleagues at Kansas State University, particularly members of the Dodds Lab, James Guinnip, Sophie Higgs, Molly Fisher, Lane Lundeen, Abu Raihan, and Shaun Baughman, for endless support, help, troubleshooting, snacks, and proofreading. Sarah Hacker and Tari Philips keep the Division of Biology running, and really appreciate how much they look out for and take care of me and every student. Additionally, I want to thank the other graduate students on the MACRO project - Bolortsetseg Erdenee, Emily Arsenault, Battsengel Dashdorj, John Costello, Solongo Altangerel, Jake Lutchen, Khaliun Sanchir, Amarbat Otgonganbat, and Mario Minder - for spending six months in the field with me, and for building a network across

continents and disciplines that let me learn more than I ever could within one lab. Thanks to Maggie Adams for editing and optimizing the big and small. I had a lot of field help from Kansas to Mongolia, and thank Lane Lundeen, Sammi Grieger, Chú Gomez, Taylor Hall, Kari Snelding, and Andrew Newcomb for making this work possible. I wish everyone a partner as supportive and deeply good as Dan Sullins – thank you, Jerri, and Ida, for everything.

Above all, I thank my family for supporting me through this process. No one in my extended family had gotten a PhD or set foot in Kansas before I moved here, but they're better versed now in limnology and bigger fans of Kansas than any of us could have anticipated. Mom, Dad, Jake - thank you for giving me every opportunity, for getting on the small plane, and for your love and care.

This research was conducted on the unceded lands of many native tribes that historically and presently live in the US Great Plains. In particular, I have lived for the past five years on the land of the Kickapoo, Kaw / Kansa, Osage, and Oglala Sioux tribes, and am indebted to their stewardship of this land and for all that I have learned while here. I am grateful for the collaboration with native students and colleges, and for the deeper understanding of our natural world.

I am extremely lucky to have spent so much time studying a topic that I can confidently see so much value and future in, and am grateful to have worked towards improving metabolism estimation and understanding. This research was financially supported by the Kansas State University, its Biology Graduate Student Association, as well as the National Science Foundation, namely the NSF-MACRO and MAPS projects.

Dedication

I dedicate this work to my grandmothers, Florence Schechner and Eleanor Davidov, who instilled in me a thirst for knowledge and love of water.

Preface

The contents of this dissertation are the product of collaboration with my major advisor, Walter Dodds, as well as coauthors from other institutions and disciplines. Chapter 2 is formatted for Limnology & Oceanography: Methods with coauthors Walter Dodds, Flavia Tromboni, Sudeep Chandra, and Alain Maasri. Chapter 3 is formatted for Freshwater Science with coauthors Walter Dodds, Alain Maasri, John Costello, Scott Kenner, Sudeep Chandra, Battsengel Dashdorj, and Flavia Tromboni. Chapter 4 is formatted for Global Change Biology with coauthors Walter Dodds and Flavia Tromboni. All authors have given consent for chapters to be reproduced within this dissertation.

Chapter 1 - Introduction

Definition

Aquatic Ecosystem Metabolism (hereafter metabolism) refers to the rates of carbon fixation (Gross Primary Production, GPP) and biological carbon oxidation (Ecosystem Respiration, ER). Their balance (Net Ecosystem Production) represents available carbon and therefore energy pathways (Dodds et al. 2019), while their ratio indicates the relative magnitude of auto- or heterotrophy. Metabolism controls both the fixation of CO_2 into organic matter as well as the oxidized flux of organic matter back to CO_2 (Demars et al. 2011; Brett et al. 2017), and, as a measure of total biotic activity, forms the basis of the food web (Riley & Dodds 2012) and encompasses primary production through decomposition, integrating over drained area (Levi & McIntyre, 2020).

History

The availability of low-cost optical sensors has enabled long-term metabolism study, though focused primarily in smaller streams in the northern temperate United States and Europe where most limnological research has been conducted (Chowanski et al. 2020; Dodds et al. 2013; Hall et al. 2016). Existing long-term and publicly available monitoring data, including dissolved oxygen, temperature, and light, have been tapped for aggregation and modeling. These monitoring data can be subject to location biases as they are often taken in larger, perennial rivers of drinking water interest. Additionally, unclear and inconsistent QA/QC may bias results based on such monitoring. Understanding the effects and relationships to metabolism is limited by data context (Hotchkiss et al. 2018).

Concepts

Metabolism is central to carbon and nutrient cycles, and thus food web function, and is responsible for disproportionately large carbon emissions relative to the surface area of freshwaters (Welti et al. 2017). The quantity of CO_2 emitted from rivers is an understudied and underestimated component of the carbon cycle despite the significance of oversaturation of CO_2 produced by respiration providing a large emission to the atmosphere (Reiman & Xu, 2018). GPP and ER also offer a relatively low-resource functional indicator of a basic ecosystem process that allows us to quantify energy patterns in our waters and drained watersheds, and how they are changing with time and anthropogenic influence.

With respect to the broadest predictive scale, the Freshwater Biome Gradient Concept (Dodds et al. 2019) suggests specific ecosystem processes in rivers can be controlled in part by the surrounding terrestrial biome. The work that I relied on most heavily in this dissertation to put my results into a broader conceptual framework was conducted at the watershed scale. This included the River Continuum Concept (RCC, Vannote et al. 1980), which describes flowing waters as having a predictable dynamic equilibrium related to predictable changing ecosystem processes over their length. The main indicator of ecosystem function in the RCC is the ratio of GPP:ER. This ratio is predicted to first increase with river size as relative canopy cover and allochthonous matter decrease allowing greater in-stream GPP, and then decrease as downstream turbidity inhibits light penetration. More abiotically rooted, the Riverine Ecosystem Synthesis (Thorp et al. 2008) posits lotic systems as hierarchically structured sections primarily defined by hydrogeomorphology, but locally variable with vegetation and therefore climate, and has been supported through study of macroinvertebrate and fish assemblages (e.g., Maasri et al. 2021). At the watershed scale, the Network Dynamics Hypothesis (Benda et al. 2004) contrasts with the RCC in viewing lotic systems as a hierarchical discontinuum, where confluences and

disturbances are primary organizing features. This interacts with the Natural Flow Regime (Poff et al. 1997) that posits temporal and spatial patterns of flow are master variables in streams and rivers. All of these build on the conceptual work of Hynes (1975), The Stream and its Valley. These concepts are not isolated pillars of stream ecology; tenets of each are present in the others, and can be variably applicable or dominant in a particular system, providing context for our results. Here I assess how well these frameworks apply to observed patterns of river metabolism.

Methods

In this dissertation, I focus on a series of measurements made in comparable ecoregions in Mongolia and the United States. I used the open channel diel O_2 method to estimate metabolism of O_2 as proportional to C via an established respiratory quotient (Demars et al. 2015). I employed the BASE model to estimate rates of GPP, ER, and K (a constant describing gas exchange rates) simultaneously in a Bayesian framework (Grace et al. 2015). The large scale and breadth of our project meant I only had a couple of days at each site, and often many sites going at once, which limited deployment length but allowed me to sample a much broader range of river sites and conditions than most have the chance to study, giving me a unique perspective on river heterogeneity and the many small decisions at any field site, and how methods may bias our findings.

Other authors have demonstrated that site heterogeneity is tied to heterogeneity of metabolism. Siders et al. (2017) linked substrate heterogeneity and thermal stratification to variation in rate estimates of metabolism. Irwin et al. (2020) and Dodds et al. (2013) found differences in rates between main channels and side or ephemeral channels, and noted how applying findings from a particular sub-habitat would bias estimates if used to represent the whole system. Chowanski et al. (2020) called for additional measurement through time and

space to better estimate metabolism across those gradients, building on Demars et al.'s (2011) work supporting the averaging of many sensors at a given reach to better account for variation. This variability is accepted, yet long-term monitoring efforts still have variable and unclear QA/QC (Zimmer et al. 2020). My data allowed direct assessment of the broader applicability of much of this earlier work. This assessment forms the basis of my Chapter 2.

Known Controls and Global Patterns

We don't have a firm understanding of the broad drivers of stream metabolism (Pastor et al. 2017) despite efforts worldwide. Primary controls vary by scale and location, among other things, but major variables for our study systems and other rivers include temperature (Jankowski & Schindler 2019), hydrology (Poff et al. 1997; Donnell & Hotchkiss 2019), light (Dodds et al. 2019), vegetation in the form of canopy (Gucker & Pusch 2006; Preiner et al. 2020) and in-stream (Dodds et al. 2004) in addition to its removal through grazing (Sarneel et al. 2014) alongside other forms of land use (Bernot et al. 2010; Hamid et al. 2020).

Through synthesis efforts by Dodds et al. (2019), Appling et al. (2018), and Bernot et al. (2010) among others, macro patterns and gradients (e.g., multiple large watersheds across continents) have been explored, including along the interrelated axes of temperature, latitude, altitude, and light. We used these global observations of metabolism to evaluate how they scaled down both to our study sites in general and how they held up in Mongolia, where metabolism data are extremely limited in the literature.

Mongolia has been shaped by pastoralism since 1300 BCE (Chen et al. 2018) but is highly vulnerable to and experiencing rapid change. Grasslands, and the Mongolian grassland in particular, have the least protection of any biome, and the greatest proportion of area that has been converted for human use from its natural state (Surenkhorloo et al. 2021). Mongolian air

temperatures are increasing at twice the global average rate (Dorjsuren et al. 2018) and the country is subject to dramatic land cover change associated with urbanization, mining, and agriculture. Crop production is increasing because of increased temperature supporting crops that could not survive previously (Jordan et al. 2018). Mining has already caused heavy metal contamination and a dramatic decline in biodiversity, wetlands, and riparian areas (Kida et al. 2018; Surenkhorloo et al. 2021). Simultaneous deregulation and urbanization patterns have led to overgrazing.

The United States Great Plains have also been shaped by grazing and agriculture, exacerbated by drought (Silcock & Fensham, 2019). This region is highly heterogeneous, with variable hydrologic conditions (Willet et al. 2018) reflected in much greater variability in river metabolism than in Mongolian sites. Monitoring and protection of rivers in the United States often focus primarily on human-related response variables - namely drinking water quality and flood risk - which do not always overlap with priorities associated with biodiversity or other disciplines (Kovach et al. 2019).

I assess the role of scale related to river metabolism in Chapter 3. I use a mixed modeling approach to evaluate most explanatory variables of metabolism at a range of scales and subsets in order to identify both the scaling potential of different models as well as to identify the most explanatory data types and scales. This was accomplished in the context of a comprehensive and interdisciplinary dataset including hydrogeomorphic data at the reach to basin scale, land use data, and anthropogenic impact, among other variables.

Relationship with Climate Change

Temperature is a primary control of river metabolism in forms ranging from direct (Metabolic Theory of Ecology, Jankowski & Schindler 2019) to indirect (eg., through

evapotranspiration's effects on hydrology or how ice prevents light penetration, Dodds et al. 2019) and the sensitivity of metabolism to temperature changes is modulated by other controls (including physical factors and nutrient limitation, Pearce et al. 2020). Respiration is expected to change more rapidly than GPP under a warming climate (Demars et al. 2011) though this has proven to be variably accurate (Song et al. 2018).

Identifying patterns of drivers and constraints of metabolism will improve our ability to anticipate how rivers (and ecosystem services therein) change alongside climate and other human impacts (Bortolotti et al. 2019). I explore both the applicability of larger scale modeling approaches to ecoregion-scale metabolism, and the utility of using such models to project change with temperature in Chapter 4. This is particularly important because of the rapid temperature increases in Mongolia paired with recent strong anthropogenic influence.

Usage and Challenges

River metabolism is a functional indicator and integrator, reflecting the quantities and fluxes of carbon as a currency across hierarchical scales and disciplines. We can use it to quantify large-scale fluxes, predict responses to global change up the food web, and track management and recovery (Ruegg et al. 2020; Arroita et al. 2019).

My research ran up against uncertainties and difficulties in metabolism estimation. Some of these were confronted directly in Chapter 2. Still, we have trouble estimating reaeration and therefore respiration when reaeration is particularly low or high. We lack low-resource ways to partition autotrophic and heterotrophic respiration. We have much less understanding of aquatic than terrestrial ecosystems despite their holding disproportionately much of earth's biodiversity. There are still major gaps in the data outside of the global north that prohibit evaluation of proposed models and controls. Different disciplines approach monitoring, evaluation, and river

study differently enough to make comparison and integration challenging. Both through papers in this dissertation and others I have contributed to while at KSU, we have addressed some of these problems and made progress towards more comparable, informed, and useful aquatic science knowledge.

Chapter 2 - How Do Methodological Choices Influence Estimation of River Metabolism?

Abstract

River metabolism modeled from diurnal dissolved oxygen (DO) has become a widely used metric of ecosystem function, yet many papers provide insufficient methodological detail for replication. Only 79% of 43 sampled papers published from 2015-2019 mention calibration, 44% describe sensor placement, and 34% did not describe estimation approaches such that the study could be replicated. Given that spatial heterogeneity in rivers influences metabolism, and measurement sensitivities vary with sensor model, it is important to have appropriately detailed information in reported methods along with a fundamental understanding of how river heterogeneity might influence metabolism. We deployed 2-8 sensors at 92 steppe river reaches to characterize site heterogeneity, evaluating how sensor placement and type, deployment length, drift correction, data source, local versus remotely sensed data, and calibration can affect metabolism estimates. Estimates of gross primary production (GPP) and ecosystem respiration (ER) were inconsistent and unpredictable depending on deployment location within a river reach; GPP and ER rates varied up to 131% and 69% respectively across a river width and up to two orders of magnitude within a reach. DO sensor brands vary in precision and accuracy; we found even when operated within stated performance range, estimates of GPP and ER could vary by 82% and 198% respectively if not calibrated beyond factory setting, as determined using field data from a sample site. Inaccuracies from sensor drift over weeklong deployments led to an average 48% ER overestimation, and 2% GPP overestimation comparing uncorrected with corrected field data. We suggest best practices for more comparable, precise, representative, and accurate methods.

Introduction

Ecosystem metabolism is central to ecosystem function and is the basis for understanding energy flows and ecological efficiencies from local to global scale. Carbon metabolism consists of carbon fixation (Gross Primary Production, GPP) and biological carbon oxidation (Ecosystem Respiration, ER), as well as their balance (Net Ecosystem Production). These properties describing organic carbon dynamics and rates of activity delineate heterotrophic and autotrophic states in lotic waters (Demars et al. 2011; Dodds and Cole 2007) and are used by managers to assess river condition (e.g., Chowanski et al. 2020). Net ecosystem production roughly represents the CO_2 emissions or sequestration by a river where groundwater influence is minimal (Hall et al. 2016). Dissolved oxygen (DO) measurements have commonly been used to measure metabolism in aquatic systems as a proxy for carbon flux; they are less complicated than multi-species bicarbonate equilibria and are detectable against low background concentration, though they do not account for anoxic processes (Dodds and Cole 2007).

We contend many stream metabolism publications do not report methods in sufficient detail to allow replication and confident comparison among studies. We aimed to understand how equipment choice, data decisions, and instrument placement in otherwise heterogenous waters influence metabolism determinations. We surveyed methods used in metabolism studies, determined how equipment and data decisions affect resulting rate estimates, and evaluated sensor placement in temperate steppe rivers to understand their influence on metabolic calculations at the local to reach scale. Our primary goal was to quantify potential sources of bias and error from initial experimental design to the final step of reporting methods such that future studies are reproducible and better characterize river metabolism.

Advances in sensors and approaches to estimating metabolism (including iterative Bayesian methods allowing for calculation of error and fit) have allowed broader estimation of stream metabolic characteristics. The more sophisticated models currently applied over the original accounting methods (Odum 1956) are informed by additional metrics including factors influencing aeration and photosynthetically active radiation (PAR) and require deciding between direct measurement or estimation of those metrics. Metabolism estimation also requires measurements of barometric pressure, reach geometry and hydrology, and temperature. Some metrics with strong effects on rate estimates have been considered previously. For example, there is ample literature on estimating aeration, leading some to suggest there are weak relationships among estimates modeled, measured, or calculated from stream and river (hereafter river) morphology and hydrology (e.g., Riley and Dodds 2012). GPP has been linked to water velocity (Edwards and Owens 1962), and alongside ER to substrate size and variability (Cardinale et al. 2002). Other metrics might affect GPP and ER estimates but are less completely analyzed. Open access and long-term datasets are increasingly tapped for aggregation and synthesis (e.g., Rüegg et al. 2020; Hoellein et al. 2013; Bernhardt et al. 2018) though the quality of those data is not always clear or consistent, but used regardless assuming that the quantity of data outweighs any QA/QC issues with a particular small subset of measured sites.

Many data aggregations do not report key measurement conditions, potentially influencing the reproducibility and interpretation of metabolic rates and factors influencing these rates. For example, StreamPULSE (Koenig et al. 2019), a large, multi-institutional effort that aggregates long-term metabolism data, includes metadata describing reach characteristics, but not the specific location or habitat type where the sensor is placed. Such aggregations generally do not document information on length of deployment, wiper use, and calibration procedures,

which can alter precision and accuracy (Hall and Hotchkiss, 2017). We assumed these data aggregations did not document this information because it was not available, which opens the question: how often are key measurement descriptions reported in data sources? Thus, we report a systematic review of the literature as the first step of this paper and explore which characteristics might most strongly influence estimation of rates.

Research Questions

This paper arose from our experiences attempting to measure metabolism in river segments (reaches) across biomes and continents. We sought to determine if particular river habitats have consistent patterns of GPP and ER and if temporal, logistical, and spatial constraints could influence metabolism. These decisions included equipment choices and methods, temporal and spatial specifics of field deployment, and approaches to data processing. We approached the following questions: (I) Do estimations of rates of GPP and ER vary with data source including sensor type and placement? (II) How does QA/QC influence metabolism estimates? and (III) How can we conduct the most representative, accurate, and replicable metabolism field study? Answering these questions can guide the development of best management practices when quantifying metabolism estimates using DO sensors within rivers. We provide a sample workflow including relevant steps discussed in this paper in Figure 2.1.

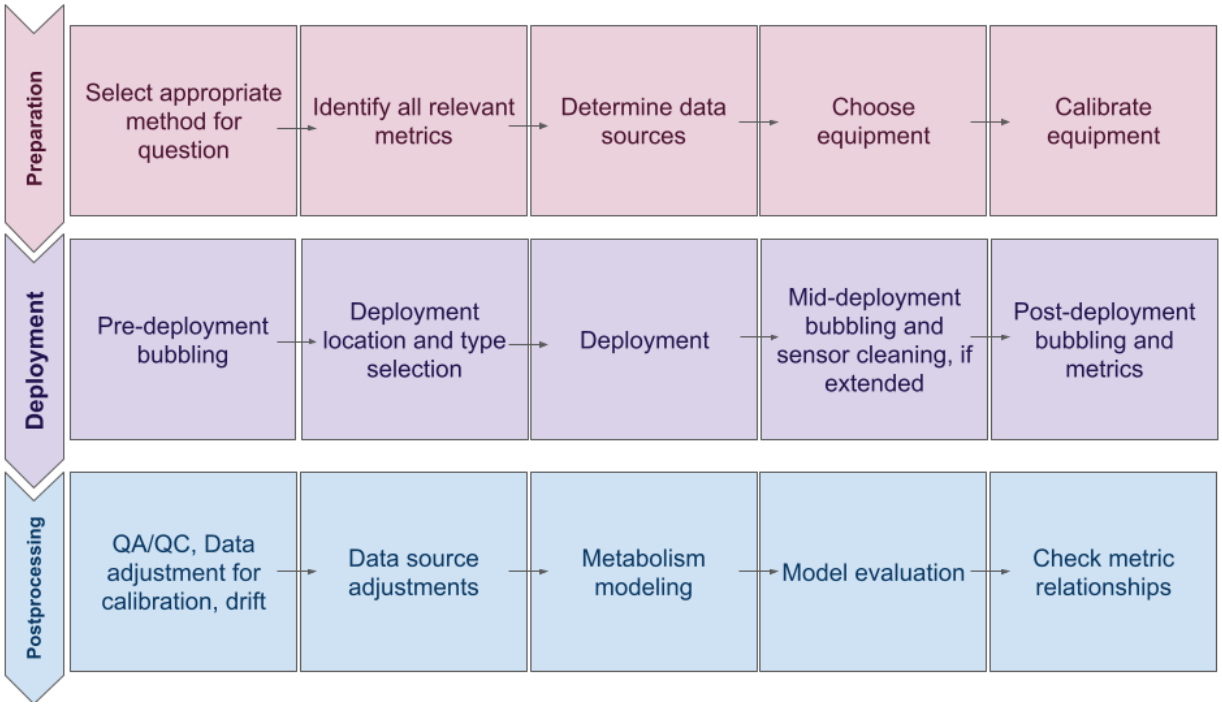


Figure 2.1. Sample workflow and decision points of note in open channel one-station metabolism estimation

We surveyed the recent metabolism literature to assess the most common practices and reporting of key metrics, and to identify potential commonalities and bias. We followed with analysis of some consequences of inaccurate or imprecise estimation of these metrics. Using field data, we evaluated how estimates can be influenced by sensor precision and accuracy, reliable sampling representative of overall river conditions (e.g. probe placement in the river), and collection in the field versus remotely obtained values of light and barometric pressure. We demonstrated that individual decision points can change metabolism estimates by orders of magnitude, and that compounded uncertainty, furthered by underreporting, prevents responsible comparison.

Materials and Procedures

Methodological Review

We examined how recent metabolism papers presented their methods, and so used a Web of Science (ISI WoS) search on 7 October 2019 with topic parameters [((river OR stream OR aquatic) AND (metabolism OR NEP) AND (diurnal OR oxygen))], refined by [CATEGORIES: (ECOLOGY OR LIMNOLOGY) AND DOCUMENT TYPES: (ARTICLE)], narrowed to years 2015-2019, and focused on the top 140 results. We discarded 96 papers for not using open channel methods, focusing on lakes or fishes, or otherwise not estimating rates of metabolism. We were interested in determining 1) what model or approach authors used to estimate rates of metabolism in rivers, 2) what was the minimum length of sensor deployment, 3) which sensors were used to monitor DO and photosynthetically active radiation, and 4) how authors described both calibration and sensor placement. Specific paper titles and information are available in Supplemental Table 2.4.

Metabolism Monitoring and Estimation

We measured DO concentration and temperature at ten-minute intervals using Precision Management Engineering miniDOT sensors (Vista, California) for periods between 24 and 144 hours, constrained by the logistics of mobile expeditions with multiple stream ecology research objectives in remote locations. At three additional sites, sensors were deployed for up to two weeks between calibrations, for a total period of fourteen months. MiniDOT files were corrected both for drift and initial calibration based on common average values during pre- and post-deployment logging together in continuously aerated water for at least 30 minutes to account for any potential change during deployment which could be attributed to biofouling or other causes of instrument drift, including slow adjustment to different conditions including temperature. Our

determination of accuracy was based on the assumption that atmospheric oxygen concentrations at each elevation are correct as are the determinations of oxygen saturation concentration dissolved in water as a function of temperature.

PAR was logged at ten-minute intervals near the site using Odyssey PAR loggers (Odyssey, Christchurch, NZ) calibrated against a LI-COR Quantum Sensor (LI-COR Biosciences, Lincoln, NE) as the manufacturer and Long et al. (2012) recommend. Placement ranged from as close as adjacent to the DO sensor to a few kilometers away at a basecamp, where particularly bare landscapes made locally deployed sensors visible and vulnerable to livestock trampling and/or theft. DO saturation as a function of temperature was calculated using barometric pressure, either measured at the site with a YSI 6020 V2 handheld unit (Yellow Springs Instruments, Yellow Springs, Ohio) or as a daily mean retrieved from the nearest Weather Underground or NOAA station and corrected for elevation. We used multiple YSI 6020 V2 handheld units that were calibrated against each other for barometric pressure and against the NOAA weather station at Manhattan KS Municipal Airport. This allowed us to check if calibration held in the field by comparing multiple calibrated instruments.

Rates of GPP, ER, and aeration were simultaneously estimated alongside standard deviations over each 24 hour period using the BASE model (v2.3, BAyesian Single-station Estimation, Grace et al. 2015) in R (R Core Team, 2013) run with 200,000 iterations and ten-minute interval data, and a theta of 1.07177. We discarded sites where we were unable to model the data with good fit as evaluated by posterior predictive check, modeled versus estimated data correlation, chain convergence, deviance, and information criteria from the model, as well as a visual evaluation of model fit. We did not estimate the relationship between aeration and

discharge in this paper as would be necessary for longer deployments in hydrologically variable rivers.

Note the BASE model output is in mass O_2 per volume per time, so results do not rely on accuracy of measurement of river hydrology and morphology (e.g. average velocity, depth, width). If we assume a DO temporal pattern is truly representative of the whole channel, then average depth upstream in the zone influencing the measurement can be used to convert the estimate to per unit area. However, when we place numerous probes in one lateral transect, we cannot know the average depth upstream of the parcel of water above each probe. Thus, our results are reported per unit volume and do not use measures of average depth, which also requires knowledge of average velocity and gas exchange to know how far upstream the measurement was influenced (Demars et al. 2011).

We calculated the upstream zone of influence as the estimated 80% turnover distance as in Hall et al. (1.61 * Velocity (m day^{-1}) / K (aeration, in day^{-1}), 2016) and report it alongside our metabolism estimates to show the sensitivity of the calculation of this distance to the aeration (and velocity) estimation. We did not evaluate or incorporate uncertainty associated with our discharge (and subsequently depth, width, and velocity) sampling methods, but saw no clear increase over our ten discharge transects. Additionally, we avoided visible lateral or groundwater inflows, which disproportionately affect respiration estimates, as explored in detail by McCutchan et al. (1998). Further, we made multiple discharge measurements along the zone of influence and did not see substantial increases in discharge that would be associated with significant groundwater input.

We used the model put forth by Riley and Dodds (2012) to estimate the initial slope of the photosynthesis-irradiance curve (α) and the maximum rate of photosynthesis (P_{max}) to evaluate

differential responses of GPP to light (Jassby and Platt 1975). The relative variation in these metrics is evaluated using the Coefficient of Variation (CV, the standard deviation divided by the mean).

Hydrology

Velocity profiles were taken at ten evenly spaced transects over at least a calculated 15-minute flow distance upstream of the DO measurement points using either a handheld flowmeter (Marsh McBirney, Hach, Loveland, CO) and topset rod at ten points per transect at 0.6 * depth, or in deeper and non-wadeable rivers using an Acoustic Doppler Velocimeter (Sontek, Xylem, San Diego, CA) pulled across each transect perpendicular to flow direction and corrected to width rather than track distance. River widths were taken at transects, and intermediate points between them, for a total of 19 locations to better characterize variability. We diagrammed site probe placement, indicating characteristics such as relative depth, location along a river width, substrate type, and other relevant details including undercut, bar, and canopy or other vegetation presence.

Study Sites

We evaluated river reaches in three ecoregions of the US and Mongolian temperate steppes in summers of 2016-2019, in addition to three locations on the Kansas River for 14 months in 2018-2019. We discuss four of these reaches in detail and include estimates from paired sensors at 23 sites (Supplemental Table 1). Discharge values ranged from 0.04 to 53 m^3/s among sites and captured a wide range of flow conditions. Mongolian rivers had generally open canopies with unstable banks accompanied by heavily grazed riparian zones, and livestock nutrient inputs. Rivers in the US more often had forested riparian zones, flow controlled by

upstream impoundments, stabilized channels, established riparian grasses, and cropland nutrient input.

We studied 98 discrete valley-scale hydrogeomorphic units across 19 river networks in temperate steppes of Mongolia and the US. These units were delineated as geomorphologically distinct using the GIS-based program RESonate (Williams et al 2013; Maasri et al. 2019) to extract valley-scale hydrogeomorphic variables from existing geospatial data. We used ten variables for this delineation extracted at 10 km sample intervals: elevation, mean annual precipitation, valley width, valley floor width (i.e., floodplain), valley width-to-valley floor width ratio, river channel sinuosity, down valley slope, geology, and left and right valley slopes. This approach ensured we had a wide range of river systems for this assessment.

We selected reaches as two riffle-pool-riffle sequences and where in-situ hydrology measurements could account for the majority of flow, therefore avoiding braided river sections with more than three parallel channels and river confluence sections. We also avoided reaches in proximity to urban areas, bridges, or other significant anthropogenic features. Reaches had at least a 15-minute travel time as calculated by a single velocity transect, with a minimum of 300 meters and a maximum of two kilometers. We deployed at least two sensors in *an area of active/representative flow* - a phrase we commonly encountered in the literature - but with intent to minimize visibility towards lowering risk of human disturbance. For example, we tied probes to large, submerged rocks or suspended them from overhanging branches. We additionally, based on probe availability at each site, sampled numerous representative or potentially overlooked but contributing "habitats" such as backwaters, debris dams, undercuts, and deep pools. In each of these multiple probe deployments, we used the same type of sensor and calibrated them together before and after deployment to minimize variance not attributable to

deployment location. Descriptions of specific placement locations by sampled reach for sites discussed in detail are in Table 1.

We also analyzed a more heavily instrumented site where sensors were placed along a horizontal and vertical transect. We placed four sensors along the surface tied to a wire, and two just above the bottom propped up on rocks, one in the thalweg and the second in an undercut (Fig. 2.8a).

Table 2.1. Characterization of sites discussed where multiple DO probes were deployed to assess spatial effects, as well as metabolism estimates and standard deviations, and the array reference column links to our figures to denote specific rates and deployment location. $O_{2,80}$ represents the 80% turnover distance as calculated in Hall et al. (2016).

	Latitude, Longitude	Altitude	Array Ref.	Placement Description	Discharge	Mean Depth	Mean Width	Atm. pressure	GPP	GPP SD	ER	ER SD	K	K SD	$O_{2,80}$
		m			m^3s^{-1}	m	m	atm	mg O_2 L^{-1} day^{-1}		mg O_2 L^{-1} day^{-1}		day^{-1}		km
Eg	50.57,	1121	A	Side, shallow	7.62	0.38	39.8	0.882	1.72	0.08	6.59	0.22	5.97	0.18	11.99
	101.53		B	Thalweg					0.91	0.03	3.75	0.08	4.02	0.07	17.78
			C	Side, constrained					0.99	0.02	3.63	0.06	3.58	0.05	19.97
			D	Side, shallow, macrophytes					27.57	8.78	252.75	74.55	26.86	7.89	2.66
				Thalweg, constrained, faster					0.85	0.02	3.15	0.07	3.69	0.07	19.38
Tensleep	44.25,	2709	A	Side, surface	2.57	0.66	10.7	0.732	0.27	0.07	5.83	0.34	5.67	0.28	8.92
	-107.22		B	Thalweg, surface					0.31	0.09	6.26	0.45	6.07	0.36	8.34
			C	Surface, btw side and thalweg					1.32	0.37	9.23	1.55	8.36	1.27	6.06
			D	Grassy side, surface					0.39	0.10	6.64	0.46	6.16	0.37	8.22
			E	Thalweg, deep					0.92	0.15	4.47	0.42	9.45	0.72	5.36
			F	Deep, undercut					0.72	0.10	7.77	0.46	6.96	0.36	7.28
			G	Thalweg, on log, surface					1.71	0.59	34.03	7.47	32.10	6.87	1.58
			H	Thalweg, deep					0.04	0.05	8.24	1.01	8.87	1.09	5.71

Delgermurun	49.64,	1284	A	Side, shallow	8.63	0.48	33.72	0.857	1.78	0.03	12.03	0.21	3.02	0.05	19.63
	99.92		B	Mid channel, deep					1.74	0.03	10.29	0.16	2.46	0.04	24.06
			C	Thalweg, deep					1.54	0.02	8.65	0.14	2.06	0.04	28.69
			D	Shoreline, slow					1.75	0.04	10.77	0.22	2.63	0.06	22.52
			E	Center, deep					1.67	0.02	9.64	0.15	2.21	0.04	26.86
			F	Side, shallow					1.79	0.03	11.79	0.16	2.60	0.04	22.79
Eg 2	50.1,	1168	A	Backwater	0.43	0.26	6.68	0.862	1.83	0.13	2.44	0.18	7.14	0.31	6.20
	101.59		B	Side, shallow					3.58	0.52	2.37	0.48	20.94	2.15	2.11
			C	Thalweg, shallow					2.06	0.26	11.49	0.94	13.92	1.01	3.18
			D	Thalweg, deep hole					2.21	0.27	15.05	1.20	13.19	0.96	3.36
Tongue	44.77, -107.47	2143	A	Btw side and thalweg	1.78	0.39	13.38	0.732	1.23	0.17	1.13	0.21	11.05	1.04	4.00

Assessment

Literature Review

We reviewed recent metabolism literature to document the amount of reported methodological detail. Our goal was to determine if the papers followed the basic scientific yardstick of allowing an independent reader to be able to replicate the measurement and subsequent estimation of rates from such measurements. Instrument calibration was mentioned in 79% of papers. Only 44% made any reference to sensor location or attachment point in the reach. About one-third (34%) of papers simply stated that rate estimates were calculated as in Odum (1956); even this classic paper provides seven possible ways of calculating aeration/diffusion. Of the papers that reported using light loggers, 41% stated they used HOBO (Onset Computer Corporation, Bourne, MA) sensors, while Long et al. (2012) showed HOBO are not cosine corrected (as is standard for estimation of sunlight flux for photosynthetic rates) and have high individual variation. Long et al. (2012) showed that light measures with these sensors could be improved by developing an exponential calibration adjustment from a LiCOR sensor and averaging output from multiple sensors. These papers did not mention using this method, though they may have done so.

Sensitivity Analysis: Equipment Choice

Equipment choice is an early decision point that could influence quality of estimates; we identified the eight most commonly used sensors in recent literature (based on the literature analysis of 43 papers described in the introduction, citations are provided in Supp. Table 2.4), with accuracies and resolutions as reported by manufacturers (Table 2.2).

Table 2.2. Manufacturer-reported accuracy and resolution of commonly used commercial DO loggers as well as the effects of individual sensor uncertainty on a sample rate from a set of field data used to estimate of GPP and ER from one 24-hour period on the Zakhvan river. Estimates are derived from adding the amount of DO reported by each company as accuracy to each reading over 24 hours, as compared to the estimates of GPP and ER from modeling the actual data based on calibrated, drift corrected probes, both from the BASE model.

Sensor	Reported Accuracy between 0-8 mg/L	Resolution	GPP (mg O_2 L^{-1} day^{-1}, as compared to 1.74)	ER (mg O_2 L^{-1} day^{-1}, as compared to 10.29)
Campbell Oxyguard	0.2 mg/L	0.2 mg/L	1.77	10.23
Driesen + Kern Logger	0.05%	1%	1.75	10.10
Hach Hydrolab	0.2 mg/L	0.01 mg/L	1.77	10.23
Hach Lange	0.1 mg/L	0.10 %	1.85	9.89
HOBO Logger	0.2 mg/L	0.02 mg/L	1.77	10.23
PME MiniDOT	5%	0.05 µmol/L	1.79	10.14
YSI ProODO	0.1 mg/L	0.01 mg/L	1.85	9.89
Orion Oxygen Probes	2%	0.1 mg/L	1.79	10.14

We used a sensitivity analysis to identify how this range of accuracies, up to 5%, might result in different estimates. We found that adding +/- 0.1, 0.2, and 0.4 mg/L (1.3%, 2.7%, 5.3%

difference relative to saturation at the site we used to make this calculation) to each DO reading over 24 hours was responsible for a maximum 82% difference in GPP but a 198% difference in ER (Fig. 2.2) as compared to the calibrated and drift-corrected trace we collected in the field from a relatively metabolically active site. To be explicit, we use percentage difference among separate points to refer to the absolute value of their difference over their average, rather than the equation for percentage change in one point over time, the difference between the final and initial value over the initial value. This illustrates how two probes, both within factory calibration but not calibrated against supersaturated water and/or each other, deployed in the same location could result in substantial differences in estimates of GPP and ER. Our tests assume sensors have similar precision while correcting for drift more directly addresses issues related to accuracy. Repeated measures in the form of multiple sensors should additionally improve precision, but calibration promotes accuracy. We picked this site on the Tongue river intentionally to illustrate the effects of high aeration and reasonably high GPP. It was, therefore, added to sites discussed in greater detail in this text, and is present in Table 2.1.

These data make it clear that relying upon factory calibration alone can lead to more inaccurate estimates than those obtained with data generated with careful field calibration procedures, and ER rates may be more strongly influenced by poor calibration than GPP. This probably occurred because GPP is estimated from diurnal changes in DO coupled with departure from saturation, whereas ER is estimated solely by departure from DO saturation.

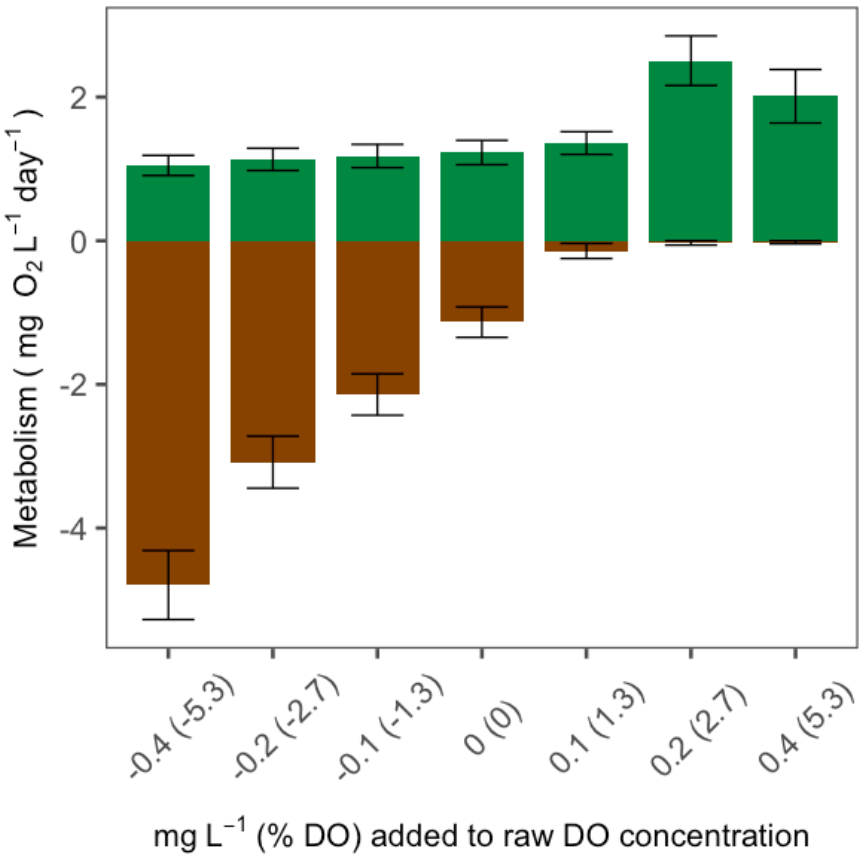


Figure 2.2. Rate estimate (±SD) variation when each DO reading is adjusted by adding up to 0.4 mg L^{-1} (5.3 %), based on one 24-hour period on the Tongue River. Specific values available in Supplemental Table 2.3.

Spatial Sensor Placement Affected Metabolism Estimates

The placement of sensors in river channels affected the estimates derived from models for GPP and ER. In particular, one site showed how deployment locations distributed vertically and horizontally across a river cross-section can yield widely varying estimates (Fig. 2.3): over 36 hours, four sensors cycled between 8.9-9.9 mg/L daily, while a fifth sensor (Fig. 2.3 Sensor D) placed in a flowing macrophyte-dominated fine sediment area cycled into hypoxia daily, despite being located within 1m of another sensor that had a minimum oxygen concentration just below 9 mg/L (Fig. 2.3 Sensor A). The other sensors varied substantially from the median

calculated while disregarding the macrophyte site, with estimates of GPP differing (Bayesian mean ± standard deviation not overlapping, shown in Table 2.1) between all but two sensors.

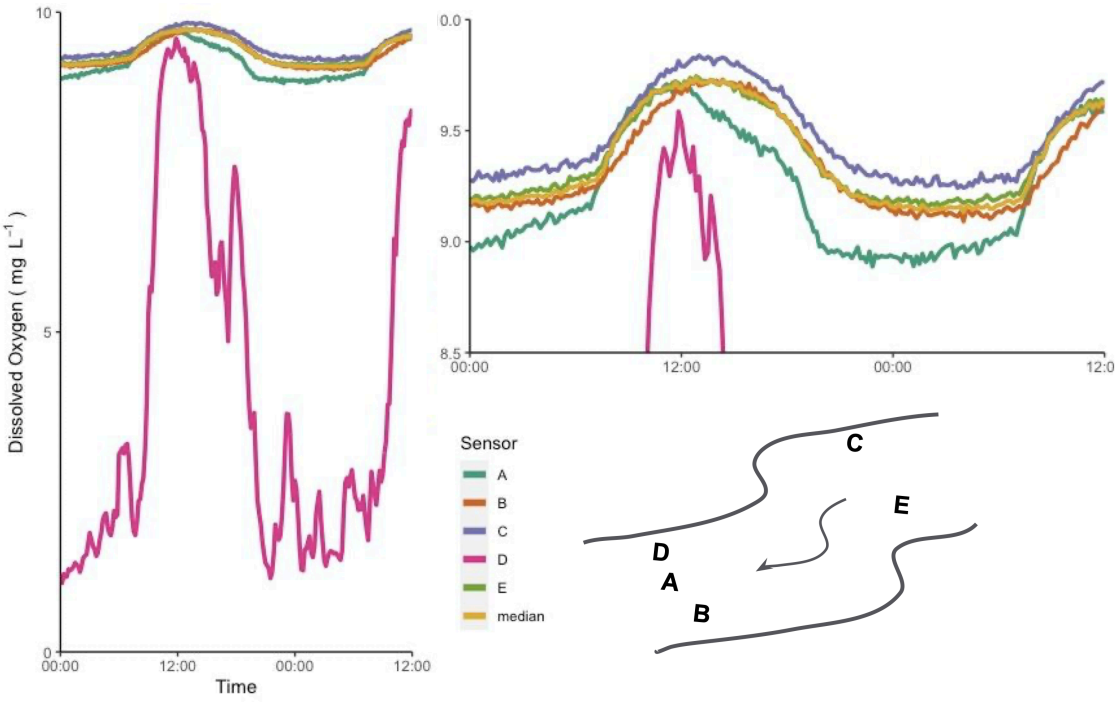


Figure 2.3. DO concentrations over 36 hours on the Eg river plotted full-scale and zoomed in for five sensors A-E as well as the median of sensors A,B,C,E, and a diagram of their arrangement in the reach. Average reach width was 39.8 m and distance between sensor sets was 575 m.

The GPP functional characteristics calculated from diurnal DO curves as in Riley and Dodds (2012) exhibited variability in estimates of **α** (the slope of the initial response to light) and the maximum photosynthetic rate (P_{max}) for each sensor at two sites (Fig. 2.4). P_{max} and **α** estimates were comparable at all sensor locations at Delgermurun (CV 0.13 and 0.35 respectively) and Tensleep (1.09 and 0.27 respectively). Thus, relying on single-point measures to calculate these photosynthetic parameters may be more reasonable than for metabolism rate estimates.

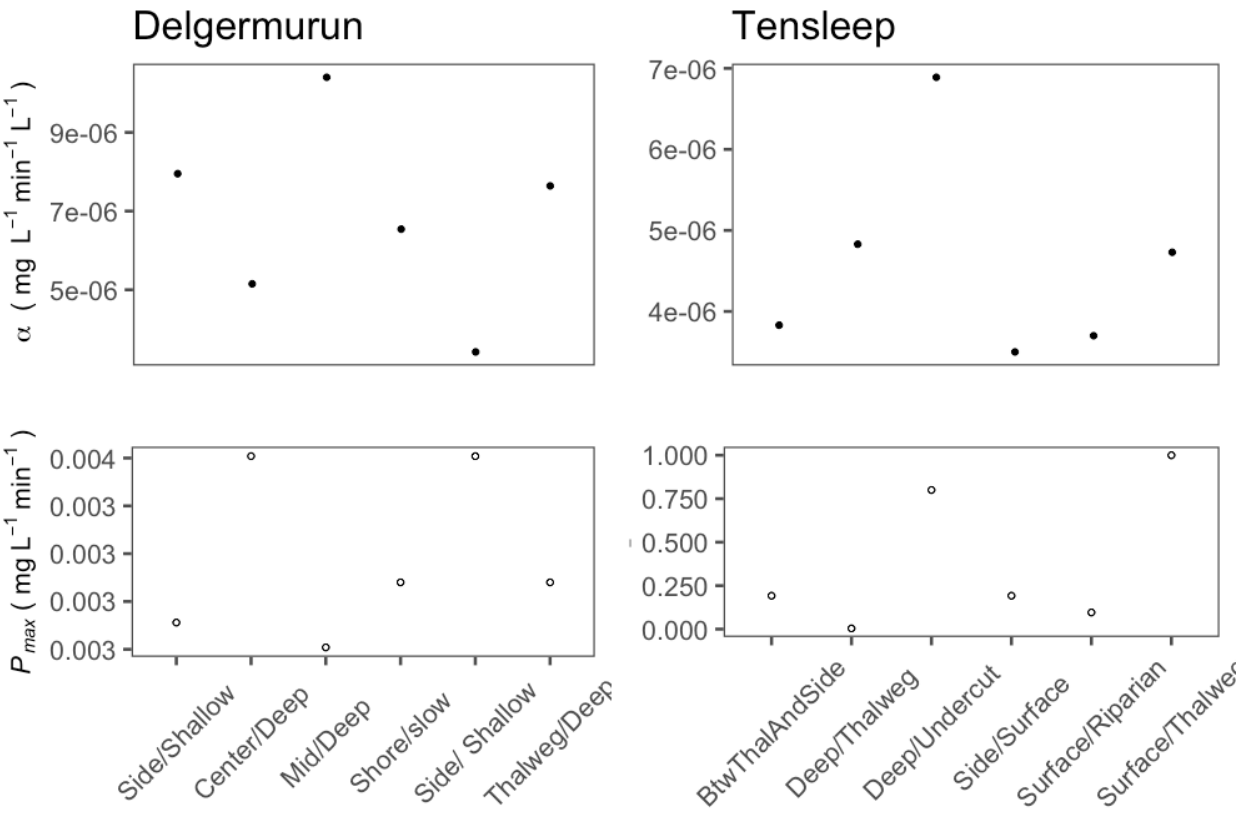


Figure 2.4. Rates of α (initial slope of response to light) and Pmax for each sensor at two sites calculated for one 24-hour period. Note differing axis ranges. CV for Pmax and α for Tensleep were 1.09 and 0.27 respectively, and were 0.13 and 0.35 for Delgermurun.

Absolute and relative ER values varied more than GPP at two of three sample sites (Fig. 2.5, CV GPP versus ER, 0.63 versus 0.24 at Tensleep, 0.05 versus 0.12 at Delgermurun, and 1.85 versus 2.06 at Eg). Inconsistent differences appeared when estimates were aggregated by broad characterization as side or center and shallow or deep (Fig. 2.6). For example, the sensor described in Delgermurun as "Thalweg/Deep" is present in both "center" and "deep" categories.

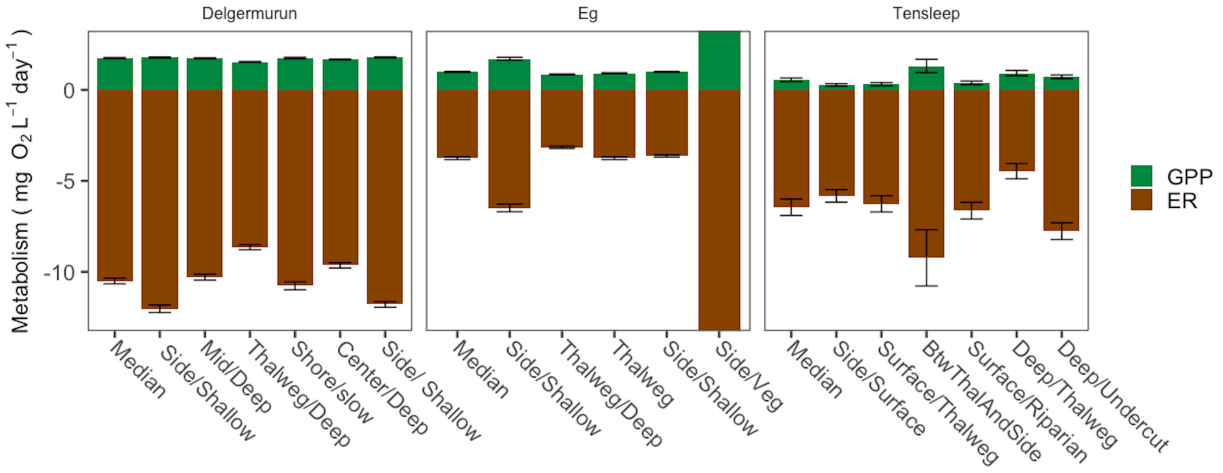


Figure 2.5. Volumetric rates of metabolism (±SD) at different locations in one area along each river length, as well as their median. The Tensleep site is a direct array (all sensors attached to a cable running across a river width), while the Delgermurun and Eg sensors are at multiple locations within a 15-minute travel time reach. The Eg sensor labeled by the "Side/Veg" bar has dramatically greater rates of production and respiration (GPP 27.57 ± 8.78, ER -252.75 ± 74.55), and is shown in more detail in Figure 2.3.

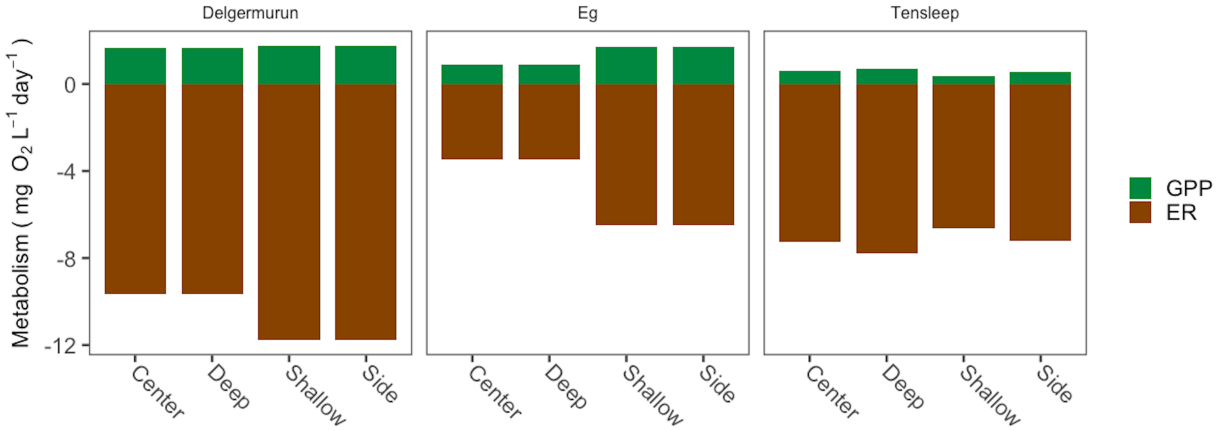


Figure 2.6. Median rates by habitat designation as side or center and shallow or deep at three sites. For example, all sites within a river in Figure 2.4 that have Side as part of the location description on the x-axis are included in the "Side" median bar.

Sensor placement was important in many sites (Fig. 2.7). Each end of each line in this figure connects GPP and ER estimates from one probe at a given site with that of a second probe in the same cross section. For example, the line could show a contrast between one shallow and

one deep sensor, or one side and one center sensor within the channel. Locations with greater ER were also generally those with greater GPP, but ER rates were greater in magnitude and variability: the median difference in GPP between two paired sensors was 0.72 g O_2 L^{-1} day^{-1} (SD 4.51), while the median difference in ER was 2.23 g O_2 L^{-1} day^{-1} (SD 9.54). Site information for these sensors are detailed in Supplemental Table 2.1.

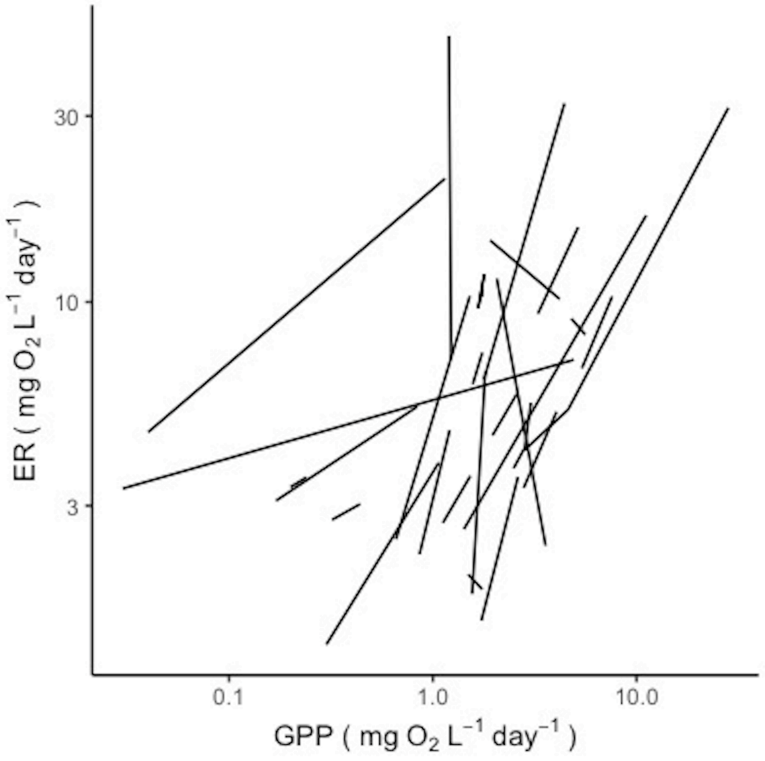


Figure 2.7. GPP and ER estimates for each of two paired sensors at a given location (37 pairs at 23 sites, shallow versus deep or side versus center, note log scale). Each line connects estimates from two separate probes deployed in the same river transect.

While diurnal DO trends appear similar at a highly instrumented transect (Fig. 2.8c), rate estimates (Fig. 2.8b) were sensitive to the apparently minor differences in the diurnal DO (maximum DO varied less than 1% but GPP and ER varied up to 131% and 69% respectively). We did not anticipate that B and C would have such disparate estimates, but we believe this difference is due to the fact that B was located in the thalweg and related to C having the biggest

error associated with its estimate. We were interested in the fact that sensors B and F also had highest P_{max} despite having the least similar, or rather most exceptional, placements. This, alongside the spread evident in some instances represented in Figure 2.7, supports Demars' (2015) recommendation that multiple sensors be averaged even in well mixed areas to better incorporate localized heterogeneity in single station estimates.

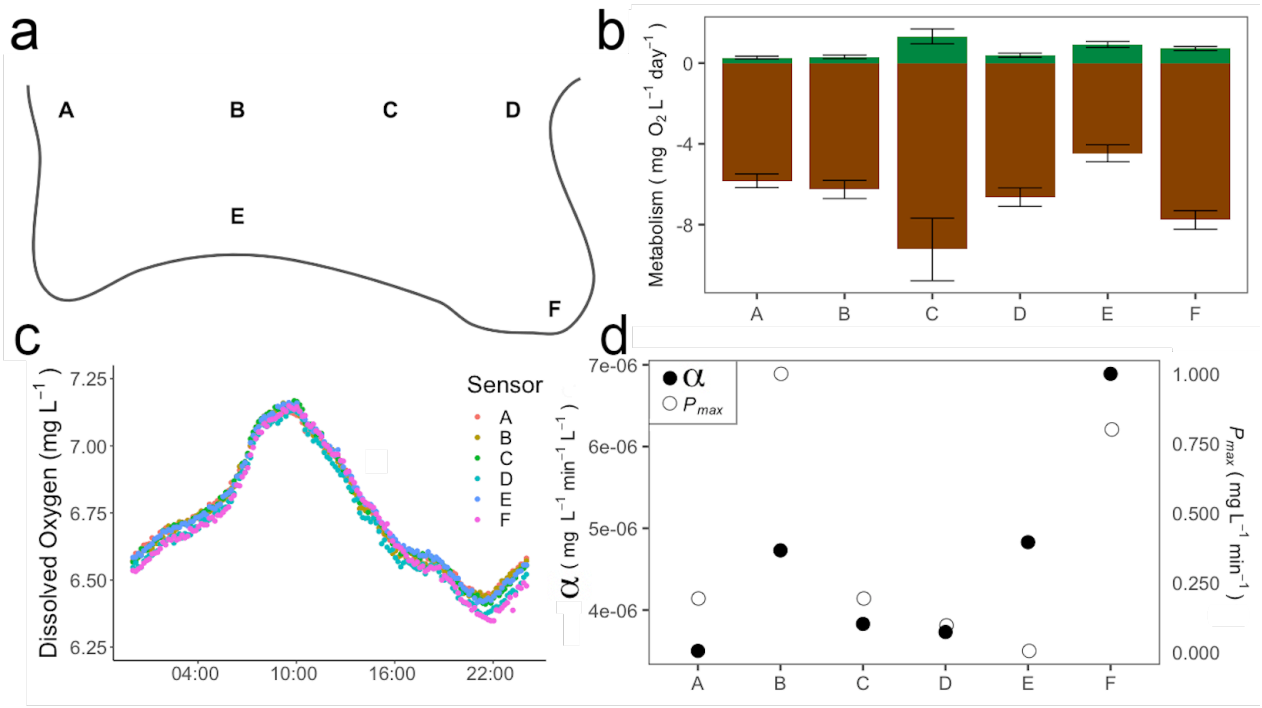


Figure 2.8. Cross section sensor array diagram where capital letters represent individual sensors (a), 24-hour rates of GPP and ER (±SD) for each sensor (b), overlaid DO concentration over 24 hours (c), and α, P_{max} for each sensor (d) all at one site and location along on the Tensleep River.

Refining Calibration Procedures

We ran all sensors together in DO-saturated water to identify sensors that were clear outliers to minimize bias by equipment as is common in other QA/QC protocols. This procedure occasionally identified sensors that had varied widely from factory calibration or that were malfunctioning. Such sensors generally could not be calibrated properly. We observed that even

with calibration, sensors can drift following deployment (related to, for example, biofouling or physical changes in the optical dyes used in the sensors over time). Our sensors were not fitted with wipers which could have decreased biofouling, and as such were not suitable for long deployment. We compared estimates for three weeks of data from one year at the same site (Fig. 2.9) with correction based on pre- and post-deployment bubbling to the same data without drift correction. Drift-corrected mean ER varied from uncorrected by up to 100% and GPP by up to 4%. While biofouling may explain drift, we cannot rule out other potential sources, and calibration based on before/after readings corrects for these as well, and should be considered for probes with and without wipers.

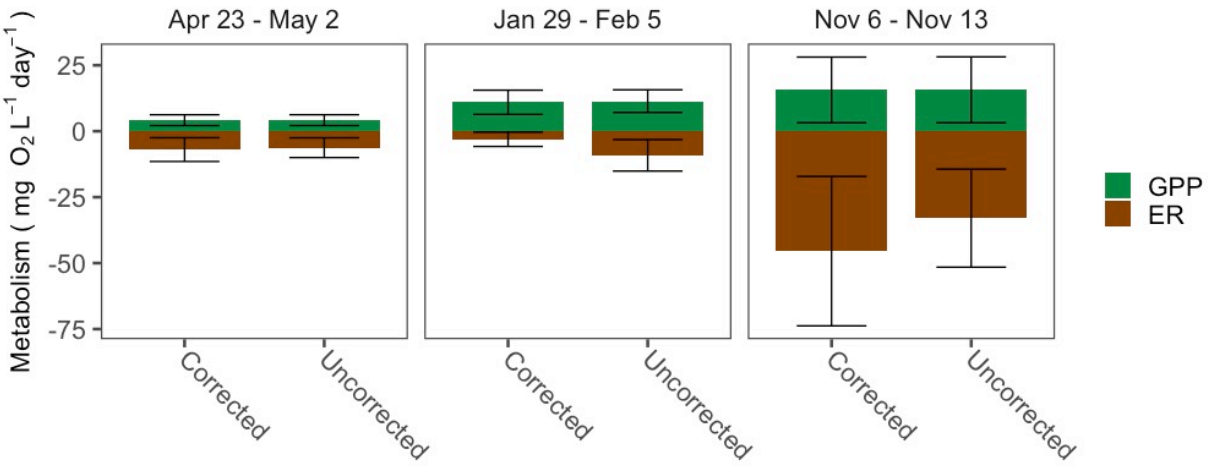


Figure 2.9. Average rates of GPP and ER (±SD) for three separate full weeks of data in three different seasons of continuous monitoring on the Kansas River near Manhattan, KS, with and without drift correction based on pre- and post-deployment runs in oxygen-saturated water.

Barometric Pressure Source Can Affect Estimates

Here we assessed differences among potential barometric pressure values. We compared a value collected at the site using a ProODO (Yellow Springs Instruments, Ohio, USA) handheld unit with barometric pressure sensor, as well as hourly average, daily average, and monthly

average from the nearest weather station with historical data as corrected to site altitude (NOAA, Worland, WY 68 km away). We also used the daily average at the NOAA site and as provided at sea level to evaluate the importance of adjusting to site altitude (2709 m, Fig. 2.10). As the saturation value of the atmosphere is a function of barometric pressure, not accounting for daily variation of atmospheric pressure could alter results based on changing influx or efflux rates of sub- and supersaturated DO, respectively.

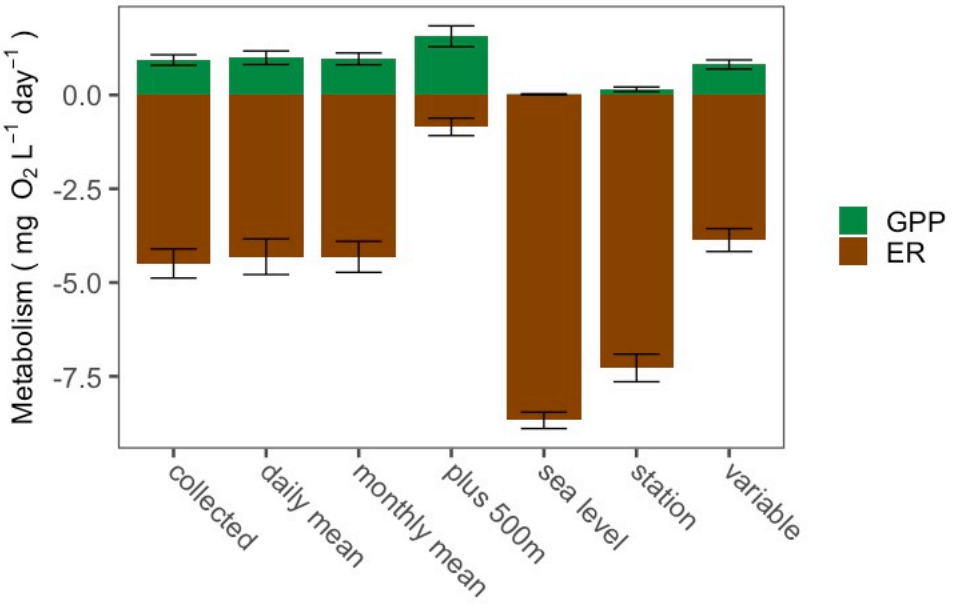


Figure 2.10. Estimates of GPP and ER for one 24-hour period as vary by barometric pressure source: based on a single handheld sonde measure at the Tensleep site in the Bighorn National Forest, WY (collected), daily mean barometric pressure as obtained from NOAA at the Worland, WY airport 68 kilometers away and corrected to site elevation (2709m), monthly elevation-corrected mean (NOAA monthly average corrected to site elevation), using hourly variable elevation-corrected data obtained from NOAA, daily mean as would be miscalculated by adding 500m to the site elevation (plus 500m), daily mean sea level uncorrected to altitude, and daily mean from the NOAA station in Worland based on station altitude (1239m). Specific parameters available in Supplemental Table 2.2.

In general, estimates should be based on continuous barometric pressure data, but in a typical day, a single value may suffice if weather patterns do not lead to strong swings in

barometric pressure. The variable data estimated lower rates of both GPP and ER compared to the daily mean of the same data. Altitude correction is essential, and even a 500 m difference can have a large impact on measurements.

Light Measurement

We note that some of the uncalibrated light probes commonly used do not provide accurate estimates of PAR, though they may have responses to light that are directly correlated with PAR values from cosine-corrected sensors. These data can be used to link diurnal DO traces to GPP by linking changes in light to rates of change in DO. However, if the parameters describing functional relationships of GPP to light (α and P_{max}) are to be investigated and reported such that others can use the estimates, they should be based on calibrated measures of PAR. Bookkeeping models (e.g., Odum 1956) would be less affected by this difference as only sunrise and sunset times are relevant, though these times could still deviate from spatially derived values based on local shading. Local shading would not be a problem on an open landscape with modest topographic relief. Models including the BLAM (Julian et al. 2008) can incorporate topographic shading alongside a range of hydrogeomorphic variables to estimate light at the water surface or at depth, but require much additional effort and only averaged accurate within 39% over more than a week of use.

After losing several PAR sensors to theft, we looked to alternatives including placing sensors at nearby and more protected basecamps, as well as by calculating diurnal PAR using geographic location, as included in and recommended by the StreamMetabolizer package (Appling et al. 2018) This model is widely used (including by Judd et al. 2009). Modeling light assumes that the location of the sensor is in an unobstructed reach on a clear day, as clouds can dramatically change PAR. When we compared the difference between measured and calculated

PAR, we had considerably different estimates for both GPP and ER at one site (Delgermurun 1.74±0.03 versus 2.10±0.02, -10.29±0.16 versus -12.64±0.10 respectively), and smaller differences at another site (Eg 0.91±0.03 versus 0.95±0.02, -3.75±0.08 versus -3.79±0.07). Total daily irradiance for Delgermurun was measured at 32 mol m^{-2} day^{-1} but calculated by StreamMetabolizer at 52 mol m^{-2} day^{-1}, while at Eg was measured at 53 mol m^{-2} day^{-1} but calculated at 44 mol m^{-2} day^{-1}. Both sites were in relatively flat, open areas, so shading from canopy cover or topography cannot explain this variability. Calculated light curves will miss these interfering factors (e.g., cloud cover, canopy, topography, or other shadows) affecting both sensors and rivers (Fig. 2.11) that can change estimates. Part of the ability to calculate GPP can be based on DO responses to these shorter-term light fluctuations.

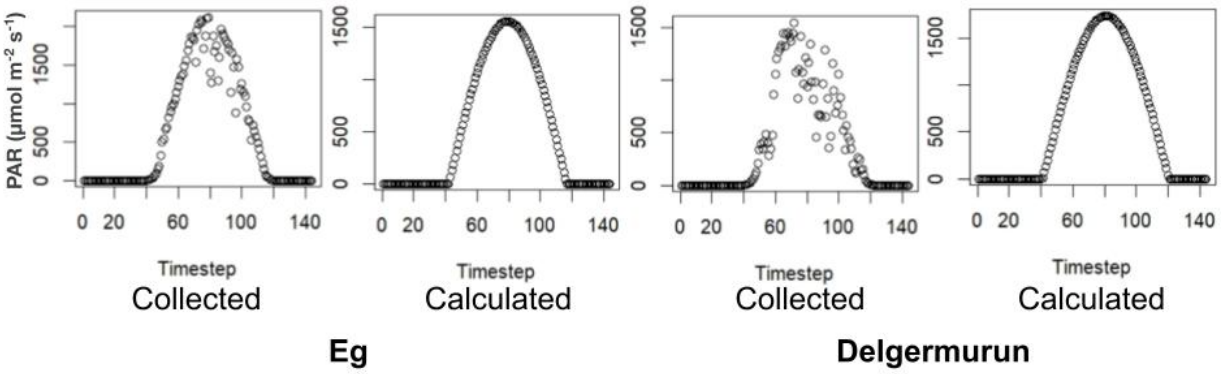


Figure 2.11. Collected and calculated PAR at two sites over 24 hours, where each timestep represents one reading from each ten-minute interval.

Discussion

We show evidence of methodological bias and underreporting in our on comparative studies of metabolism estimates and the literature, but by no means have provided an exhaustive examination of each methodological decision point. We highlight some of the decisions to be

considered and attempt to prioritize methodological practices most likely to reflect reality, given site and resource constraints.

We show that stream system heterogeneity can influence metabolism measures. Heterogeneity has become better appreciated by lotic ecologists as the discipline has matured (Fausch et al. 2002, Frissell et al. 1986), but is still not often directly addressed in whole-river metabolism studies, though heterogeneity on scales from biofilm assemblage to the river continuum has been documented (Cardinale et al. 2002) and may be used (e.g., by incorporating lateral and subchannel inflows) to better assess aquatic-terrestrial linkages and watershed context (Demars 2019). Demars' (2015) equations to calculate the percentage turnover of DO typically indicate 80% gas turnover rates in hundreds of meters for small streams and in kilometers for rivers. Our data suggest DO measured in well-mixed, main flow areas is most likely to provide results averaging across more upstream heterogeneity, though some sites on the sides or bottoms of main channels can deviate substantially from areas of main flow. Demars (2019) specifically averaged multiple diurnal curves to account for such heterogeneity, noting that this propagates additional uncertainty from each individual sensor.

Sensor deployment in a location of intermediate depth is also important as thermal stratification can cause the diurnal oxygen cycle to deviate from the main channel and can interfere with obtaining good model fit. This is particularly problematic if the sensor is below the thermocline but also could be important downstream as stratification breaks. Such stratification may be responsible for some of the wide differences in Figure 2.7, as sensors were placed in relatively deep or bankside locations. We found evidence of daily thermal stratification of pools in at least one of our sites, and as such would avoid pools for sensor placement, also discussed in Siders et al. (2017). Additionally, we did not quantify the effects of groundwater inflow (though

we saw no visible increases in discharge over the ten measured transects). Groundwater influx could be part of the reason we saw different oxygen dynamics at a shallow versus deep sensor at the same thalweg location, consideration of groundwater evaluation as discussed in Hall and Tank (2005) seems prudent.

Including the additional sources of uncertainty we examined in modeling and estimation could be done in the form of additional priors to Bayesian models, or by adding the uncertainty ranges associated with heterogeneity to bookkeeping approaches, improved by Monte Carlo simulations as in Demars (2019). The additional error from any particular metric may seem insignificant relative to the error internalized in our models. We showed sample sites not chosen to represent extreme conditions. Site heterogeneity aside, the sensitivity analysis representing sensor calibration (GPP and ER showing a maximum 82% and 198% difference respectively), barometric pressure source given correct altitude (20%, 15% difference), light source (19%, 20% difference), and drift correction (4%, 100% difference) demonstrate how this error can compound quickly (summed percentage difference 125%, 333%).

The mathematics behind metabolism calculations from DO from a single station measurement assume homogeneity in the channel. In practice, river biogeochemists assume monitoring of 1-2 (usually 1) locations averages all areas and metabolically relevant actors (Hall and Hotchkiss 2017). However, several different scales of heterogeneity may interfere with such averaging. Reichert et al. (2009) and Dodds et al. (2018) document substantial, multiscale metabolic heterogeneity, both examining data from serial reaches and offering empirical approaches to demonstrating heterogeneity and calculating appropriate reach lengths based on upstream influence distance. These data in aggregate suggest careful determination of what constitutes a “representative” reach is required to obtain results reflecting general metabolic rates

in a river. Few studies we are aware of do this in addition to Reichert et al. (2009), Demars et al. (2015), and Dodds et al. (2018).

Comments and Recommendations

We found a number of decision points that influenced metabolism potentially leading to variable estimates. The largest differences in resultant rate estimates in this assessment were associated with deployment location and accounting for sensor accuracy and drift, though α and P_{max} were more variable within than among sites. Less important were differences associated with saturation calculations (barometric pressure value as long as altitude correction was employed) and categorical channel position (shallow versus deep, side versus center designation). ER was more sensitive to most methodological choices than was GPP. McCutchan et al. (1998) found that the greater sensitivity of ER decreased at higher magnitudes of GPP and ER, and reflects larger uncertainty in ER than in GPP in streams with lower rates of each. This result is likely because GPP is driven by diurnal variation and is less affected by aeration uncertainty, while ER estimates are derived directly from the exact difference of DO from saturating concentrations.

We found sensors placed in different areas of active flow gave different rate estimates of metabolism. Sensors placed in the thalweg but off the bottom gave the closest to mean rates. Equipment choice clearly influenced outcomes, improved further by careful calibration and QA/QC procedures. Finally, locally measured versus remotely sensed light and barometric pressure resulted in different rate estimates, to a lesser degree assuming low topographic relief and accurate altitude.

Calibration of all sensors used for metabolism estimates is important. However, given that many papers do not report calibration protocols, it is difficult to know how to assess the

reliability of estimates presented in those papers. Our data show that calibration of DO probes can be one of the most important factors influencing estimates of metabolism.

Increasing reporting of methods will increase the value and utility of the data. Our review found incomplete methods reporting of sensor preparation, data QA/QC, deployment location, modeling approach, and parameter sources. Any of these differences in methods would have altered metabolism estimates, some substantially. Without including this basic information in papers, data-harvesting initiatives, monitoring networks, and management decision making, our analysis suggests that the possibility to repeat the measurements is not being met by a considerable portion of the literature. This information is also fundamental for comparative and meta-analysis.

Our recommendations are generally simple enough to adopt, and we empirically show that following general operational guidelines and reporting can improve the value and precise comparison of estimates among studies. In order of priority with respect to measurement methods, 1) Carefully calibrate all sensors before and after deployment (particularly DO sensors), and correct for drift based on calibration before and after the period of measurement. Do not rely on factory "calibrations"; 2) Deploy sensors in or as close as possible to the thalweg. Ensure that sensors are not placed along/are oriented away from the bottom or sides of the channel and not placed in areas with poor mixing with the rest of the channel; 3) Whenever possible use supporting data (light, barometric pressure) taken in the river or as near as possible, and logged at the same frequency as the DO; 4) pay particular attention to the fact that barometric pressure may be reported corrected to sea level; 5) report all measurement and calculation approaches, including calibration procedures, probe placement, data cleaning steps, and programs used for estimation. The percentage error associated with each of these steps may

vary with site conditions. We omit many points of consideration that would be crucial for longer-term deployment, which would be best served by a preliminary study evaluating different habitats and upstream zone of influence.

Next Steps

Finer spatial scale data could provide a more complete accounting of DO flux across an entire cross-section of a river. This estimate could be accomplished by deploying arrays of calibrated DO sensors in tandem with data from an Acoustic Doppler Velocimeter yielding discharge estimates associated with each DO measurement point. This would allow complete accounting of DO flux for each timepoint, and avoid problems of representativeness. This could also aid in the process of linking the contribution of small habitat differences to reach-scale production and respiration, and in identification of greater or lesser need for increased sampling intensity. This approach would be costly, require substantial effort, and only give an estimate for one cross-section of a river system.

A much broader, more detailed, and comprehensive modeling effort may specifically quantify all known error based on synthetic data and a range of possible physical factors including those not discussed here. This effort could be based on observed ranges of aeration, GPP, ER, barometric pressure, temperature, and light variability. This type of sensitivity analysis could more specifically rank the potential biases we document in this current paper, and create response curves of sensitivity of estimations that could be used to refine equipment choices, approaches to sampling and measurement, and data analysis.

Acknowledgements

We profusely thank an anonymous reviewer, Dr. Scott Ensign, and Dr. Benoit Demars who reviewed and much improved an earlier version of this manuscript. We also wish to thank

Khaliun Sanchir, Lane Lundeen, Molly Fisher, Samantha Grieger, Andrew Newcombe, Jesús Gomez, Bonnie Teglas, Kari Snelding, Emily Ulrich, and Taylor Hall for field contributions. This research was supported by National Science Foundation grants (NSF Macrosystems 1442562 and 1442544 and NSF EPSCOR **OIA-1656006**). We are aware of no conflicts of interest.

Chapter 3 - Controls on Steppe River Metabolism Vary by Scale and Network Location

Abstract

River metabolism likely has linear and non-linear controls decoupled from global gradients affecting terrestrial energetics. We expand the geographic extent of metabolism estimates derived from dynamic dissolved oxygen measurements by adding data from 89 sites in three corresponding ecoregions across the temperate steppe of Mongolia and the United States. We link estimates of metabolism with reach-to-watershed-scale metrics representing hydrogeomorphology, vegetation, climate, and anthropogenic impact to evaluate predictors and applicability of traditional ecological frameworks in the Anthropocene. We expected that metrics based on vegetation and climate related to ecoregion would have greater explanatory power than human or hydrogeomorphic data. We present the most explanatory structures of variables by river type, scale, and location. This required a systematic approach to identify the most explanatory variables, many of which were strongly correlated. We conclude that macrosystem-scale studies require broader interdisciplinary and multi-scale assessment for prediction and capture of variation in aquatic metabolism, and that observed distributions of spatial patterns of river metabolism depend on the scale of interest. This suggests that universal models explaining factors controlling river metabolism will not perform as well as those built on the scale being studied or managed.

Introduction

River metabolism estimates are a key indicator of ecosystem state in rivers worldwide (Jankowski et al. 2021). Rates of carbon fixation (Gross Primary Production, GPP) and

biological carbon oxidation (Ecosystem Respiration, ER), as well as their balance, (Net Ecosystem Production, NEP) relate aerobic biotic activity and integrity to water quality and basic ecosystem structure, and integrate ecosystem processes over the upstream watershed (Riley & Dodds, 2012; Dodds 2013; Levi & McIntyre 2020). GPP and ER are major controls on riverine organic carbon fluxes (Demars et al. 2011). They transport, transform, and release carbon, with a disproportionate contribution to climate regulation and the global carbon budget relative to their area (Cole et al. 2007). Thus, understanding scaling of processes that control river metabolism, and how those factors vary with location, will assist our understanding of river ecosystem functions across broad climatic gradients (Dodds et al. 2018).

Rivers are teleconnected (Heffernan et al. 2014) systems embedded in a socio-environmental matrix with numerous potential human influences. A macrosystem approach is therefore necessary to extrapolate existing mostly reach-scale measures in developed countries into areas that are not well studied, to account for ecosystem-scale processes, and to predict future changes in river networks (Dodds et al. 2021; Tromboni et al. 2021). Existing synthesis of river metabolism has been limited by datasets, disciplines, and level of interconnectivity considered (Hotchkiss et al. 2018). GPP and ER have nonlinear relationships to upstream conditions (Feijó-Lima et al. 2018, Feijó-Lima et al. 2019). Metabolism measurements are taken at the reach scale, so inference at larger scales is still challenging (Pastor et al. 2017) and limited by data availability in less studied areas. Upscaling is further complicated by seasonality and a/biotic river characteristics (Gucker & Pusch 2006).

Here we compare and contrast rivers in the United States and Mongolia, classified as having watersheds in similar temperate steppe ecoregions, including dry terminal basins in the western portions of these countries, and montane, and grassland ecoregions moving eastward

(Olson et al. 2001). Both countries have had long histories of human habitation. While Mongolia's landscape is characterized by intense pastoral grazing and associated instream nutrient and riparian impacts dating to 1300 BCE (Chen et al. 2018), the US has had more recent mechanized agriculture, damming, channelization, and water abstraction. These rivers are subject to three critical "trajectories of change" in metabolism of river biomes (Bernhardt et al. 2017) - rising temperatures, land use changes, and regulated flow. These three shifts are interrelated - for example, as warming allows more crops to be grown in Mongolia, stimulating land conversion, land use and flow may change as well.

We present here a large metabolism dataset for rivers in Mongolia, with matched sites in three ecoregions of the US, placed in a broad and interdisciplinary context of variables including climate, hydrogeomorphology, land use, vegetation, and human influence data at reach to country scales (Fig. 3.1). Specifically, we evaluate a) what are the rates of river metabolism in temperate steppe Mongolia and the US, b) what variables and data types best predict metabolism at each site, and c) how do various scales of data aggregation alter these predictions? We hypothesized that human development and land use would strongly influence river metabolism overall, but that hydrogeomorphology and network location (both in terms of local functional processing zones and patterns along the river continuum) would be particularly important in the climate-sensitive temperate steppe (Vannote et al. 1980; Thorp et al. 2008). We expected differences in metabolism by country and ecoregion, driven by country-specific climate, land use, and human influence. National land use practices and patterns of development were predicted to be stronger drivers than natural differences. We anticipated that metabolism data would be primarily explained by the hydrogeomorphic suite of characteristics collected - hydrology and geomorphology, in addition to adjacent land use and impact.

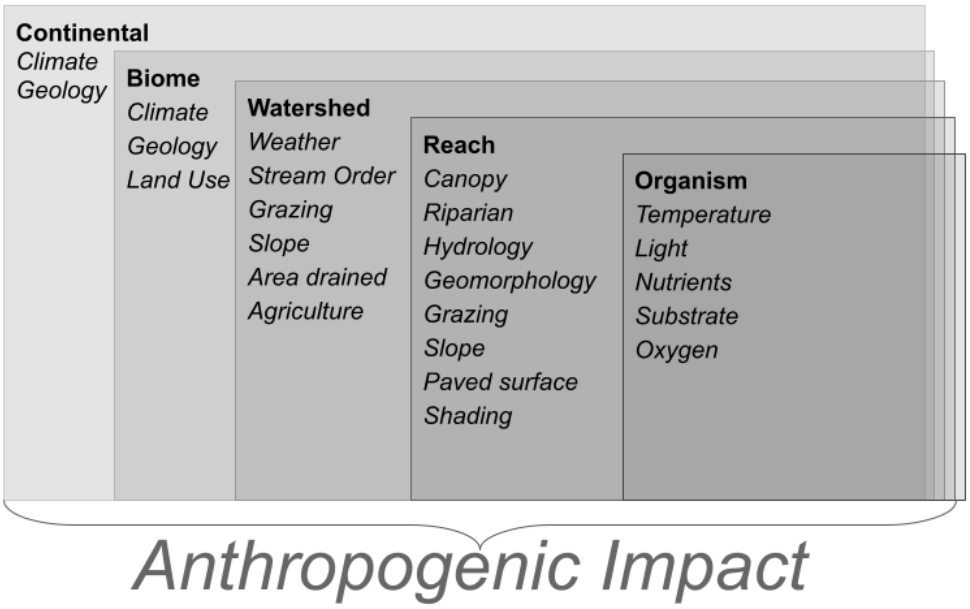


Figure 3.1. Conceptual diagram outlining relevant scales and sample variables, with influence of anthropogenic impact at all levels

Our work uses this approach across many sites, and using the same team, equipment, methodology, and control for factors that we have documented to bias estimates (Chapter 2). Levins (1966) suggested tradeoffs between generality, precision, and realism where only two of these can be achieved simultaneously, suggesting generality of broadscale synthesis sacrifices either precision or realism. Thus, we obtained multi-disciplinary data including extremely detailed and matched datasets of hydrogeomorphology up through watershed-scale remotely sensed land use information. This allowed us to maximize realism in determining factors that potentially controlled metabolism. Our sites were also intentionally selected by the RESonate model (Thorp et al. 2008) to be representative replicates of different hydrogeomorphic patches (Functional Process Zones, FPZs) in multiple watersheds, allowing us to assess the importance of reach scale properties that could directly influence the physiological controls most often

collected alongside metabolism data. We use this data context to evaluate how we can better predict metabolism at multiple scales and in the context of the metabolism literature.

Materials and Methods

Study Sites, Reach Definition

We studied three temperate steppe ecoregions (Terminal basin, TB; Montane steppe, MS; and Grassland steppe, GS) as characterized by Olson (2001) in the United States and Mongolia. To capture a range of conditions across river networks within and across ecoregions, we chose sites using the RESonate framework which defined distinct hydrogeomorphic patches (Thorp et al. 2008; Maasri et al. 2019; Maasri et al. 2021). We selected sites within regions delineated by the RESonate framework 1) that had a reach of two riffle-pool-riffle sequences, 2) where we could make hydrology measurements that would likely account for the majority of flow, 3) without highly braided areas (5+ parallel channels) or areas where large portions of the reach would have fully saturated riparian areas (e.g., extensive riparian wetland), 4) with no nearby upstream urban areas, bridges, or other significant anthropogenic features. We concentrated our hydrologic and geomorphic measurements based on a minimum fifteen-minute travel time, with a maximum of two kilometers. Sites were sampled during times of low flow, increasing comparability (Levi & McIntyre 2020). Specific site characteristics, description, and probe placement information are detailed in Supplementary Table 2.1.

Site Specific Field Hydrogeomorphic Characterization

We sampled each site using Physical Habitat protocols from Environmental Monitoring and Assessment Program Section 7 (Lazorchak et al. 1998). We used recorded measurements to calculate metrics representing habitat and dominant reach geomorphic processes (Kaufmann

1999). We sampled over a total reach length of 40 times the average wetted width, except where total reaches would have exceeded 5 km, in which case we halved length. We spaced transects at 0.1 of the total reach length, with half transects at 0.05x total reach length. Riparian cover was estimated based on recorded visual estimates of the amount and type of cover provided in a 10 m by 10 m area on the left and right banks at a transect. We recorded visual estimates of the amount and type of fish cover provided in 5 m upstream and 5 m downstream area in and over the water at a transect. We collected human influence (e.g., mowed area or building) as a presence metric that also indicated closeness to the river at a transect (P- Present > 10m away, C- Present within 10m, B- present on the bank, 0- Not Present). We weighted presence data by proximity and averaged across the transects (B- 1.5, C- 1.0, P- 0.667, 0-0). Channel geometry data included five depth measurements across each transect, and wetted width at each transect and half transect. We recorded 100 thalweg depths along the reach length. Bank geometry data was collected at each transect on both banks and included top-of-bank elevations and distances, bankfull elevations and distances, and bank angles. We collected additional metrics in ArcGIS using digital elevation models and aerial photography to extract slope and sinuosity.

Field Methods and Sensor Specifics

We used the single station open channel diel O_2 method (Demars 2015). We deployed O_2 and temperature-logging miniDOT probes (PME, Vista, CA) at at least two locations at each site separated by a 15-minute travel time as calculated using initial manual velocity measurements using either a Sontek Acoustic Doppler Velocimeter (Sontek, San Diego, CA) or a topset rod and Marsh-Mcbirney Flo-Mate (McCrometer, Hemet, CA). Deployments ranged from 24 hours to a week in length. MiniDOT sensors were calibrated and adjusted for drift during deployment, and sensors placed in representative flow were averaged to give a representative site estimate

(Chapter 2). Odyssey photosynthetically active radiation sensors (PAR, Odyssey, Christchurch, NZ) were deployed as closely as possible to each site but as several were lost to theft or animal disturbance, some sensors were deployed from a nearby basecamp when river reaches had high human presence. The Odyssey sensors were calibrated seasonally against a Li-COR Photometer (Li-COR, Lincoln, NE).

Metabolism Estimation

We used the BASE (BAyesian Single Station Estimation, V2.3, Grace et al., 2015) model to obtain estimates of GPP, ER, and aeration (k) using our ten-minute interval O_2, temperature, and light data, in conjunction with average site barometric pressure. We discarded sites where we were unable to model the data evaluated by the metrics of fit associated with the model, as well as a visual evaluation of modeled versus observed data. We re-ran sites where we were unable to get a good model fit using the BASE model with the Riley & Dodds (2012) workbook method. These sites were mainly those with high gradient/aeration. Data from the site were discarded if this approach still gave poor fit based on visual observation. This left us with 89 sites where we were able to model metabolism successfully, GPP/ER was realistic, and hydrogeomorphic and watershed data was also collected successfully. For a representative value at each site, we averaged over multiple days and sensors from representative areas of flow.

Estimates of metabolism were corrected to 20ºC to account for daily temperature fluctuations and sampling times using Equation 3.1.

Equation 3.1. Temperature correction of Metabolism rate estimates following Song et al. (2018)

$$rate\ at\ 20ºC = \frac{2.523 * e^{(0.0552*20)}}{2.523 * e^{(0.0552*t_1)} * rate\ at\ t_1} \quad (1)$$

where t_i is stream temperature, using a curve built from the temperature dependence relationships of metabolism data in Song et al. (2018) and as evaluated in Riley and Dodds

(2012). We used this approach to make estimates more comparable across sites with different daily conditions, as previous work had not found a relationship between daily metabolism and water temperature (Song et al. 2018).

Data Preparation, Variable Reduction, and Evaluation

All analyses were conducted in R (v. 4.0.2, R Development Core Team 2017). Rates of metabolism were significantly nonnormal, as were common transformations (square root, log, rate/standard deviation, 1/rate, log(max(rate)+1-rate), (rate - min(rate))/(max(rate)-min(rate)), Shapiro-Wilkes all $p < 0.00002$).

The joint metabolism, hydrogeomorphology, watershed-scale, land use, and climatic datasets contained 180 variables, many of which were functionally or mathematically redundant. We built a mixed model using river as a random effect to account for spatial autocorrelation wherein sites on the same river were inheriting the same water as a sort of flow-directional legacy effect, and for all similarities within a given river that we did not include in this analysis. All variables are available in Supplementary Table 2.2. This was done over five steps to minimize bias and maximize considered variables given limited study sites (Figure 3.2). The broadest model was built (1) by selecting the top five variables based on R^2 from each individual linear regression and Spearman rank correlation with the rate estimate, and adding ecoregion and country variables where more than one was present in a subset. We use subset to refer to the sites contained in a particular model: i.e., all sites in Mongolia, or the Temperate Steppe, or below median width. If two redundant or inversely correlated variables would have been selected (e.g., percentage of fast and slow water variables which sum to one), the variable with the higher R^2 was used. If multiple forms of the same term appeared, the highest correlation value term was used (e.g., if min/max/mean precipitation or first/third quartile of sediment size would have

otherwise been selected). Additionally, where watershed land use in km^2 and in the more comparable percentage of watershed appeared, the percentage of watershed form was preferentially used. All models from each step are available in Supplementary Table 3.2.

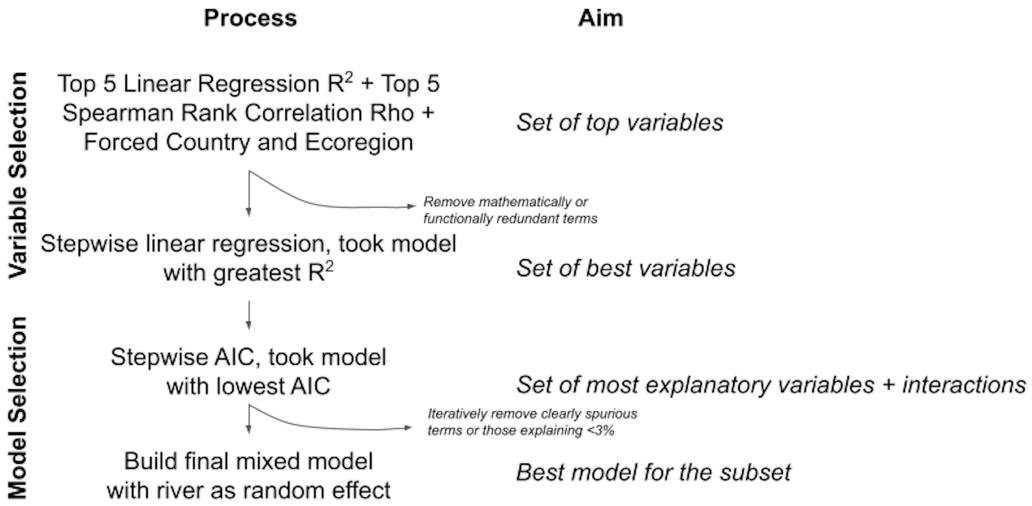


Figure 3.2. Model selection workflow, informed by Olden & Poff (2003) and Bernot et al. (2019).

Stepwise linear regression was then performed (2) to identify the best linear model from those variables using ols_step_best_subset() within *olsrr* (Hebbali, 2020). The best combination of variables was selected (3) using the highest adjusted R^2 value. We then used a stepwise Akaike Information Criterion procedure (4, stepAIC, within *MASS*) to refine the model and identify most relevant interactions. This best model was expanded into a mixed model (5) by adding the river as a random effect, and refined by dropping terms explaining less than 3% of the model as well as clearly spurious or three-way interactions. Mean differences in GPP and ER were assessed using Wilcox and Dunn tests depending on the number of groups, within the *FSA* and *stats* packages respectively (Ogle et al., 2021; R Core Team, 2020).

Results

We were interested in whether climatic differences would explain differences in metabolism by country and ecoregion, so we evaluated how the relationship with climate variables varied by subset. The temperature regime is much more homogeneous and colder in Mongolia than in the US, with annual averages below 0ºC. Temperature and precipitation are positively correlated with each other and with GPP and ER at the broadest scale, except for a nonsignificant correlation between GPP and precipitation; neither temperature nor precipitation were correlated with either GPP or ER in Mongolia. In the Terminal Basin temperature and precipitation were again positively correlated with metabolism except for GPP and precipitation, and in the Montane Steppe there were no significant correlations. Grassland Steppe sites had only significant correlation between ER and precipitation. In our wider (by wetted width, a proxy for river size) sites, GPP and precipitation were positively correlated, while ER and temperature were positively correlated for both narrow and wide (below or above median wetted width) sites, in addition to ER and precipitation and GPP and temperature being positively correlated in narrow sites. When climate and human influence variables were both present in a model, and for ER models, climate variables were more explanatory than human influence metrics. Human influence variables were not explanatory at the country scale, and climate was more explanatory for narrow than wide sites. Our results varied by scale and found that most explanatory variables did include these broader datasets and were scale dependent.

Rates of Metabolism

GPP ranged from 0.01 to 22.56 mg O_2 m^{-2} day^{-1}, while ER ranged from 0.03 to 46.55 mg O_2 m^{-2} day^{-1}. Our estimates of metabolism fall within the range of published values of GPP and ER (Supplementary Figure 3.1). Rates of GPP and ER were of lower magnitude and

variability in Mongolia as compared to the US, as is illustrated by the density distribution of rate estimates (Supplementary Figure 3.2). GPP did not vary by country alone (Wilcoxon rank sum test p=0.27) but ER did (p=0.03). GPP and ER varied by ecoregion only between the Montane Steppe and Terminal Basin ecoregions (Dunn Kruskal-Wallis multiple comparison, P adjusted via Benjamini-Hochberg =0.004 and p=0.03 respectively), though the Terminal Basin had the highest aeration and variation among sites, likely driving differences. Within Mongolia, rates did not vary by ecoregion (Dunn p values all >0.31), while in the US, GPP and ER differed between GS-TB and MS-TB, (Dunn p for GPP GS-TB 0.01, MS-TB 0.0002, for ER GS-TB 0.03, MS-TB 0.0006). Rates did not vary by country within the GS or MS ecoregions, but both GPP and ER differed between the MN and US TB (Wilcox GPP p=0.0006, ER p=0.002). Summary statistics of metabolism estimates are available in Table 3.1. Differences were almost entirely related to the unique nature of Terminal Basin sites.

Table 3.1. Summary statistics of GPP and ER estimates

		GPP				ER			
	N	Min	Median	Mean	Max.	Min	Median	Mean	Max.
ALL	87	0.01	1.58	2.6	22.56	0.03	3.53	6.38	46.55
US	44	0.02	1.77	3.19	22.56	0.03	3.53	6.38	46.55
MN	43	0.01	1.56	1.71	5.02	0.39	3.3	3.88	14.23
TB	34	0.01	3.16	4.37	22.56	0.03	5.08	10.65	46.55
MS	25	0.02	0.76	1.20	4.63	0.53	3.02	3.29	5.84
GS	31	0.04	1.50	1.78	5.02	0.06	3.39	4.19	11.84
Wide	45	0.01	1.83	2.47	10.26	0.06	3.39	4.15	14.23
Narrow	45	0.02	1.28	2.72	22.56	0.03	3.62	8.61	46.55

Broadest Model

When analyzing all sites together for variables that drove GPP and ER, climate variables (including ecoregion defined by climate) served as the primary explanatory terms, likely indicating additional differences associated with vegetation and grazing (Figure 3.3). ER was primarily explained by GPP (34%) but 40 year mean precipitation represented 35% of ER, in addition to ecoregion (22%). ER was primarily explained by watershed and larger scale variables (specifically 40 year mean precipitation, ecoregion, country, and their interactions), while more than a third of GPP was explained by reach-scale variables, in addition to climatic and human influence characteristics. Both rates were significantly influenced by cross-scale interactions.

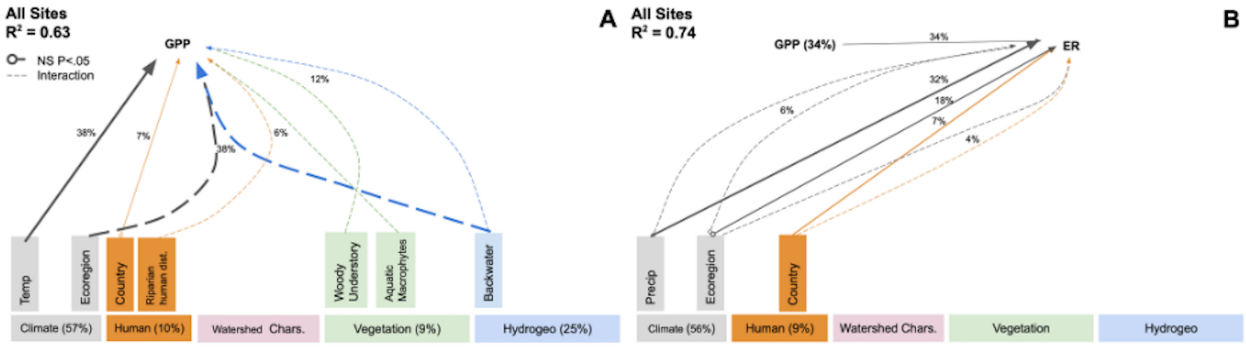


Figure 3.3. Best mixed model for GPP (A) and ER (B) across all sites. Random effect river not shown. More detailed explanation of each variable is available in Supplementary Table 3.2.

Ecoregion within Country

We didn't have enough sites to build and evaluate complex models for ecoregions within a country - 12-18 sites each. We instead list the top (absolute value) four spearman rank correlation R^2 values between our variables and estimates of GPP and ER for each ecoregion within each country (Table 3.2). These clearly, at this scale, vary by country, ecoregion, and rate. GPP and ER means may not have differed by ecoregion and country, but the relationships of

these different variables - many of which are related or otherwise different proxies for the same phenomena - varied.

Table 3.2. Top spearman rank correlation ρ of variables and estimates of GPP and ER for each ecoregion (TB=Terminal Basin, GS=Grassland Steppe, MS=Montane Steppe) subset within each country (US = United States, MN = Mongolia).

TB	US			MN			
	GPP		ER	GPP		ER	
	Ground shrubs, saplings	0.65	Cobble	0.68	40 Yr Mean Precip	0.89	Incision Ht	0.74
	Deciduous % Basin	0.63	WW x Thal Depth	0.65	Riparian Pasture	0.88	Basin elevation mean	0.67
	3L Woody Riparian	0.57	Wetland % Basin	0.65	Riparian Ag	0.88	GPP20	0.65
	Woody groundcover	0.56	Bankfull ht	0.63	Bare % Basin	-0.79	Bankfull ht	0.64
GS	US			MN			
	GPP		ER	GPP		ER	
	Riparian human inf.	0.68	Basin mean slope	-0.82	Deciduous rip. midlayer	-0.67	Min precip	-0.66
	Width	-0.68	% cover from LWD, boulders, over-hanging veg, undercuts	0.84	Rip. canopy + midlayer	-0.62	Min temp	-0.64
	Bankfull width	0.67	40 yr precip mean	0.84	Woody rip. midlayer	-0.62	Velocity	0.64
	Flood prone width	0.67	Fish conceal. exc. alg, macs (areal)	0.81	Understory woody shrubs, saplings	-0.62	Discharge	0.64
MS	US			MN			
	GPP		ER	GPP		ER	
	Q3 Substrate Size	-0.86	LWD Vol in Active Channel	-0.76	Longitude	-0.85	Basin slope SD	-0.65
	Fine Gravel % of Substrate	0.82	Undercut Bank Areal Cover	0.73	Discharge	0.83	% Riffle	-0.60
	Median Substrate Size	-0.81	LWD, % of Channel Substrate	-0.72	Flood prone width	0.80	Basin slope mean	-0.59
	Canopy Density at Bank	-0.81	Longitude	0.7	Bankfull / Depth	0.78	Riparian Canopy	-0.56

Network Location

We were interested in functional differences by river size and evaluating the relevance of the River Continuum Concept in a different landscape than the temperate deciduous forest which was used as the original basis for examples of its applicability (Vannote et al. 1980), thus we used the median wetted width (16.13 m) to split sites as wide versus narrow (Figure 3.4). This split resulted in human influence, watershed, and (mostly) vegetation variables not appearing in best ER models, but present in GPP models. ER and GPP were not significantly different by wide/narrow designation (Wilcox p=0.375 and 0.410, respectively). ER and GPP did differ marginally within MN (Wilcox p=0.048 and 0.06, respectively) and ER differed within Terminal Basin by wetted width (Wilcox p=0.030).

GPP at narrow sites was explained well by terrestrial ecoregion and related climate variables, including 40 year mean temperature (29%) and an interaction between ecoregion and backwater areal percentage. At narrower sites, ER was also primarily explained by temperature alone and as it interacts with percentage of the reach that was cobble and riparian vegetation. GPP explained 16% of ER, roughly the same as the amount explained by the three layer riparian reach portion (15.5%), and to a lesser degree percentage cobble at 9.5%.

The GPP at wide sites was most explained by percentage of shrubland at the watershed scale (53%) while ER was explained almost entirely by hydrogeomorphic variables (71%) namely the percentage of boulder/ledge area of the reach.

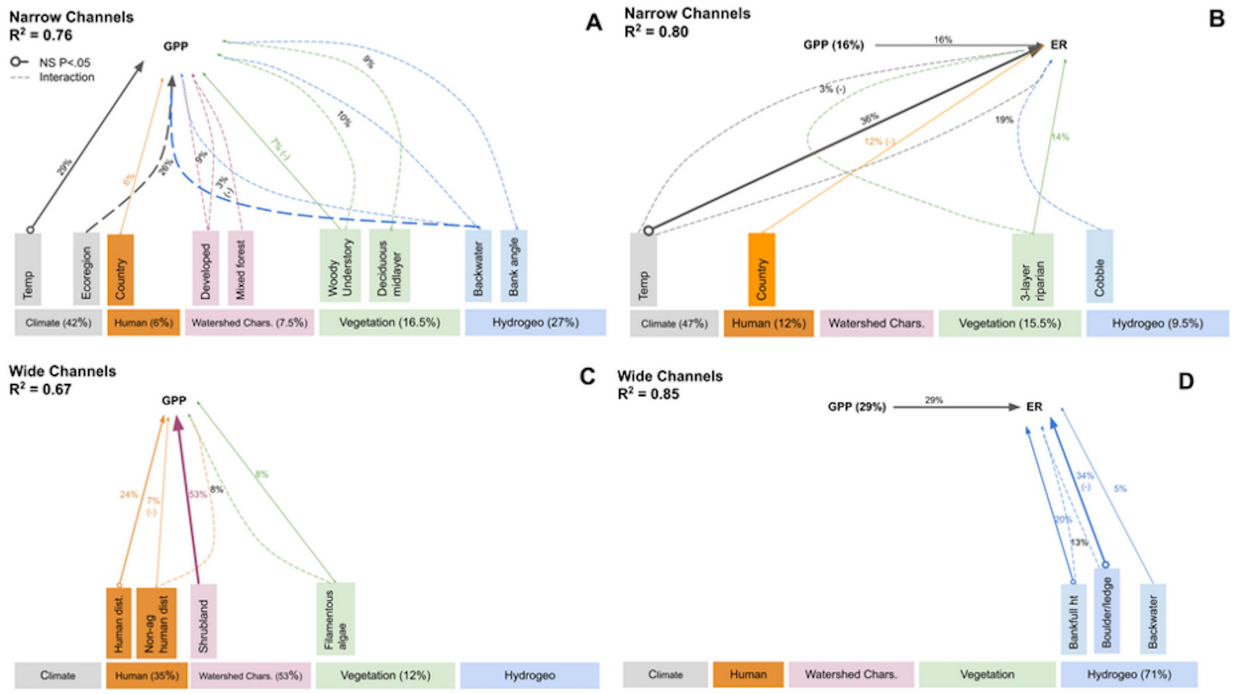


Figure 3.4. Best mixed model for GPP and ER by wetted width as narrow or wider than median. Random effect river not shown. Detailed explanation of each variable is available in Supplementary Table 3.2.

Ecoregion Scale

At the ecoregion level (Figure 3.5), the Terminal Basin had the widest range of aeration values (with rates more than twice those found in other ecoregions), but neither GPP nor ER were explained by hydrogeomorphic data or watershed-scale variables. Vegetation variables appeared in GPP models for each ecoregion. Data types and variables explaining GPP and ER did vary by ecoregion: the only commonality was that climate variables did not appear in best models for GPP in any ecoregion, but vegetation was the most explanatory data type for GPP, as climate variability was not sufficient to explain variability in the data. Regarding ER, no variable was common to all three ecoregions. The most explanatory variable of GPP was 40 year mean temperature, though backwater- ecoregion interactions were also important at 45%.

In the Terminal Basin, GPP was primarily explained by vegetation, particularly large woody debris negatively related to GPP, which we hypothesize is a proxy for shading or other indicators of allochthonous carbon (i.e large woody debris indicates large riparian trees). Additionally important (18%) was the interaction between road proximity and large woody debris. ER was primarily explained by GPP, precipitation, and country - GPP being primarily vegetation, likely signifying climate/biome constraints again.

In the Grassland Steppe, both GPP and ER, were primarily explained by vegetation. GPP was additionally explained by country and riparian human disturbance, while ER was primarily explained by substrate cover. Montane Steppe GPP and ER both were strongly related to light and factors that could influence it - GPP primarily, positively by riparian canopy and midlayer (72%) and secondarily discharge. ER was principally explained by both large tree canopy and undercut bank presence at more than half the percentage explained, followed by shrubland negatively related to respiration.

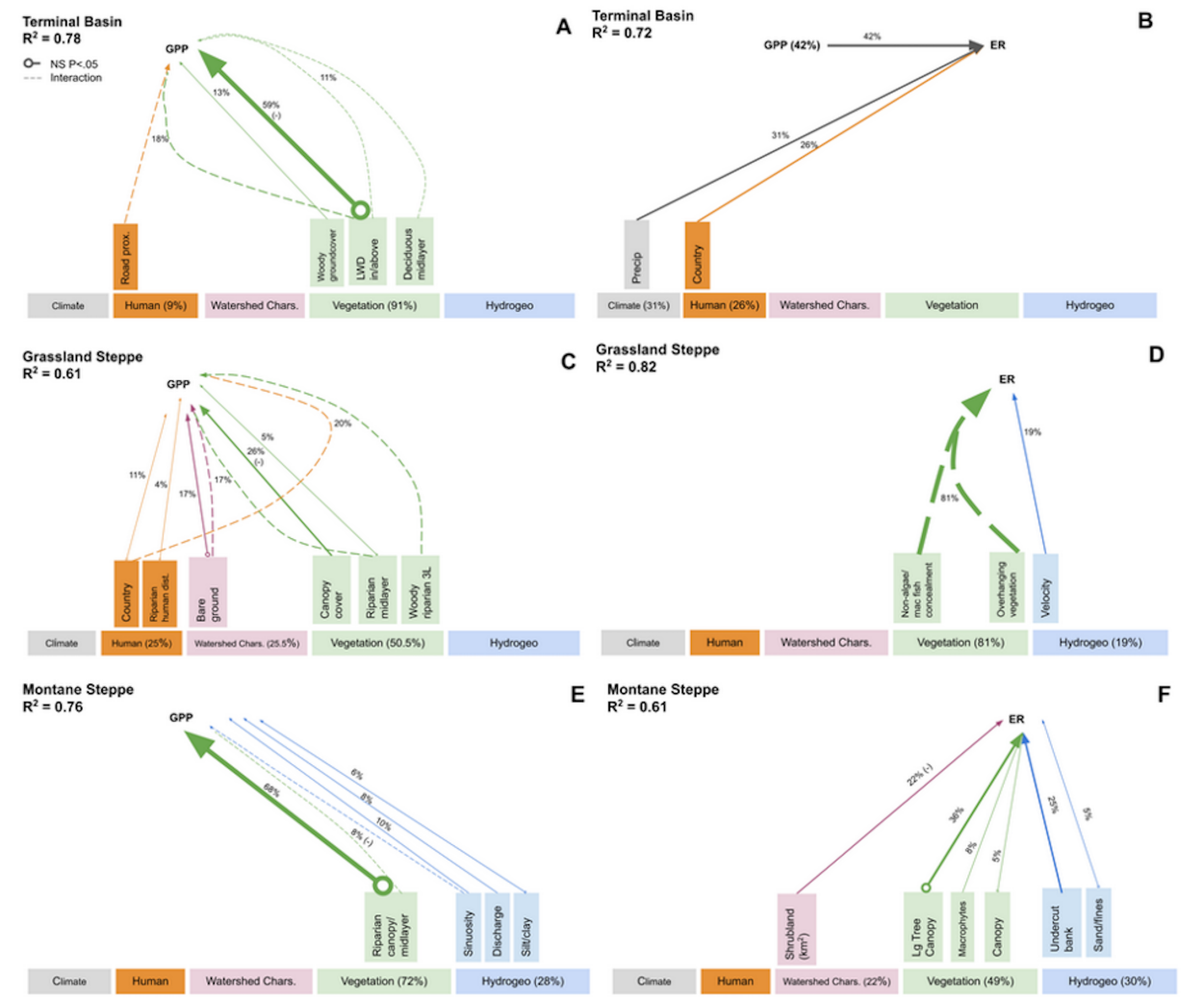


Figure 3.5. Best mixed model for GPP and ER within ecoregion. Random effect river not shown. Bottom row indicates general variable category, with specific terms above in the same color. Detailed explanation of each variable is available in Supplementary Table 3.2.

Country Scale

Climate is much less variable in Mongolia (Figure 3.6A, 3.6B) relative to the US (Figure 3.6C, 3.6D) - much lower and more tightly constrained values of temperature and precipitation caused these terms to drop out at this scale while vegetation variation - and likely grazing - represented most variability. GPP in Mongolia was dominated by vegetation and, to a lesser degree, discharge (Figure 3.6A). Riparian agriculture was also influential (12%) likely in the form of grazing pressure. ER was primarily explained by hydrogeomorphology, namely bankfull

height and boulder/rock ledge area, while vegetation variables did not appear (Figure 3.6B), potentially representing the disturbance associated with large flows. Shrubland was present as well, likely as a proxy for grazing pressure at the watershed scale.

In the US, 38% of the variation in GPP was explained by ecoregion (Figure 3.6C). Additionally, deciduous land cover at 42.5% and in interactions with cobble/velocity (19.5%), also separate sites. ER was primarily explained by GPP (at 43%) and flow proxies (slope and cobble, 36%). ER was also negatively related to deciduous canopy presence, likely as a temperature effect (Figure 3.6D). Vegetation variables appeared (though minimally) as explaining variation in US ER but not GPP in an opposite manner from MN where vegetation explained GPP but not ER.

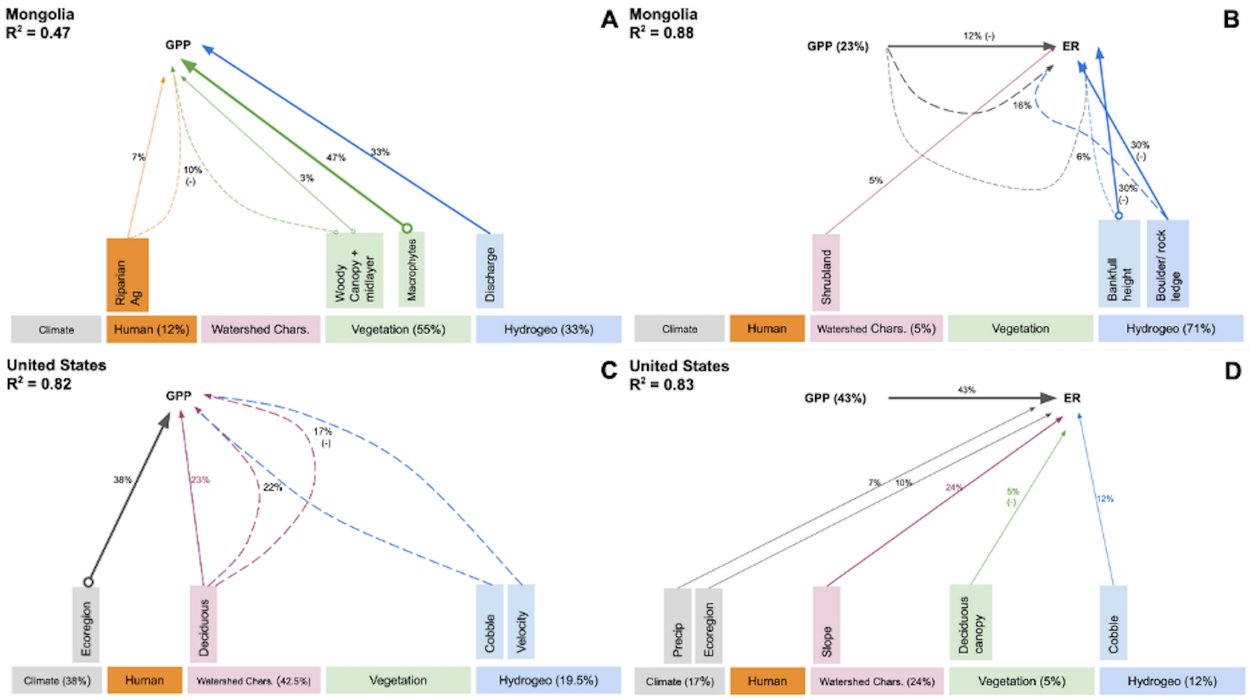


Figure 3.6. Best mixed model for GPP (Panels A and C) and ER (Panels B and D) within country (Mongolia and United States). Random effect river not shown. Bottom row indicates general variable category, with specific term above in the same color. More detailed explanation of each variable is available in Supplementary Table 3.2.

Model Sums

We additionally compared the best mixed models among our datasets, by summing the percentage contribution of each data type, and dividing interaction percentages in half (Table 3.3, represented graphically in Supplementary Figure 3.3). This allowed us to more easily see that vegetation was important to all data subsets except US GPP. At the ecoregion scale, vegetation was most explanatory, and climate did not appear. Climate and hydrology were important at narrow but not wide sites. For ER, GPP was explanatory in the best models for every subset

except for GS and MS. Human influence metrics were not important to ER at the country level, for wide sites, or in the GS and MS ecoregions. Climate and human variables were important at narrow but not wide sites, and GPP and hydrogeomorphic variables were important at the wide vs narrow scale. For both GPP and ER, climate was most important at the global scale as well as for narrow sites. Climate was the data type explaining the majority of variation in each model (Table 3.3) for both metabolism rates for all of the data together, as well as for narrow sites. Vegetation was the dominant feature for the Montane Steppe and GS ecoregions, and for MN GPP. Watershed characteristics were dominant for GPP for the Wide and US subsets. GPP was most explanatory for ER in the US and in the Terminal Basin.

Table 3.3. Percentage contributions to each best model by data type. Sums range from 99-101 due to rounding.

Subset	Rate	GPP	Climate	Human	Watershed	Vegetation	Hydrogeo
ALL	GPP		57	10	0	9	25
MN	GPP		0	12	0	55	33
US	GPP		38	0	42.5	0	19.5
TB	GPP		0	9	0	91	0
GS	GPP		0	25	25.5	50.5	0
MS	GPP		0	0	0	72	28
NARROW	GPP		42	6	7.5	16.5	27
WIDE	GPP		0	35	53	12	0
ALL	ER	34	56	9	0	0	0
MN	ER	23	0	0	5	0	71
US	ER	43	17	0	24	5	12
TB	ER	42	31	26	0	0	0
GS	ER	0	0	0	0	81	19
MS	ER	0	0	0	22	49	30
NARROW	ER	16	47	12	0	15.5	9.5
WIDE	ER	29	0	0	0	0	71

Discussion

We found that metabolism models were generally specific to the scale and scope of the data upon which they were built. Early metabolism papers sought to identify controls by looking at direct, proximal variables that constrained metabolism physiologically (e.g., whether light or nutrients were more important). More recently, limnologists have attempted to scale up these patterns and understandings by aggregating as much data as possible, and attempting to fit to one representative model to evaluate drivers and relative importance. These models are then used to try and predict metabolism based on local conditions. However, we did not find cross-scale predictability based on fundamental physiological drivers. We therefore evaluated how explanatory variables varied with scale.

Deriving Importance of Physiological Controls with Broad Synthesis

Broad scale syntheses have evaluated potential controls and drivers on metabolism, with varying scale, scope, and data context. Our finding that ER was often predicted by GPP was consistent with work by Rodriguez-Castillo et al. (2018) evaluating scaling of metabolism across multiple watersheds, additionally finding that GPP was best explained by catchment size, while ER was best predicted by cross-sectional area in addition to GPP, supporting the RCC. This is not surprising as primary producers must also respire.

Other factors have also been important in predicting metabolism. Mulholland et al. (2001) found GPP and NEP correlated with PAR but not ER. Savoy et al. (2019) found that larger watershed area better explained productivity, in addition to temperature and discharge, which were more often present in our best models. Bernot et al. (2010) compared metabolism as a function of land use across regions within the LINX II project, and found that GPP was highest

where riparian vegetation was minimal and vice versa down to lowest GPP values in reference and forest sites. ER was more variable but was controlled by temperature, organic matter, velocity, and GPP, among other variables. GPP and ER were both explained by land use category, and by important interactions: anthropogenic nutrient loading mattered more when limiting, and the response of ER to land use varied regionally. Land use categories - particularly the percentage of the watershed as shrubland, mixed forest, bare ground, and deciduous- explained much of the variation in several of our models for GPP and ER, most notably for GPP within a river width class. In contrast to our sites, Bernot et al. found that ER models were more complicated than GPP models, and that GPP wasn't explained by PAR in impacted sites. Bernot et al. (2010) additionally added to the evidence that one and two station estimates of metabolism did not differ significantly.

Bernhardt et al. (2018) described patterns of metabolism based on available data as annual regimes, and also found difficulty predicting reach scale patterns from broader models due to local controls. For example, global light data patterns don't translate to timing of peak light at the stream surface as a result of shading. They found that controls - including the relative importance of temperature - varied much more in aquatic than terrestrial systems. These findings are consistent with the fact that models at specific scales vary in the factors that offer the most explanatory power (or, as mentioned in the introduction, reality must be sacrificed for generality (Levins 1966)). Bernhardt et al. (2018) proposed a river "climate" consisting of hydrology, light, and temperature, which we evaluated independently at each scale in our models. Their work and our findings both support applicability of the network dynamics hypothesis through the importance of hydrology and discontinuity, the river continuum concept through the importance of relative canopy and river size, and the freshwater biome gradience concept through the

relationships with climatic controls. Hall et al. (2015) found much greater autotrophy in fourteen streams of a geographic region overlapping our sites in the United States, though these were much larger rivers on average so autotrophy would be more likely following decreased canopy without sufficient depth to preclude establishment of primary producers, following the river continuum concept.

We explain variation in steppe metabolism using interrelated variables across different scales and disciplines. We cannot interpret the dominance of grassland vegetation in explaining respiration without acknowledging its presence as constrained by larger scale temperature, precipitation, geology, or land use, and understanding that even those variables affect each other bottom-up and top-down. We discuss most explanatory variables by scale, understanding that all of them are likely latent variables of many other terms and scales; Jankowski & Schindler (2019) found that geomorphology, e.g. slope (or valley form) controls temperature sensitivity (and therefore retention, decomposition, connectivity) as mediated by organic matter production (via storage, grain size, connectivity, residence time).

Many subsets of our variables were collinear, so we were able to evaluate how different metrics of the same phenomenon were predictive of metabolism: velocity was not explanatory, but the percentage of a reach with cobble or larger substrate was which we expect to relate to areas of greater water velocity. Velocity is representative of conditions during sampling and substrate distributions are representative of bankfull (2-year return period) flooding. Thus, the main shaping processes of the river (Wolman and Miller 1960; Poff et al. 1997) operate on a long temporal scale, where typical channel changes occur.

Watershed Scale

We separated our sites by median river width to assess functional separation guided by the River Continuum Concept. Climate was more explanatory for GPP and ER in narrow sites, but its variables did not appear in best models for wide sites, reflecting a higher sensitivity of smaller systems to climatic changes (Follstad Shah et al. 2017). Our findings supported these transitions with channel size along continuua in the RCC, and as in Reisinger et al. (2015). Specifically, the variable that explained the highest percentage of each model for the narrow sites related to climate. We take this to mean that climate drives biome, and as is expected in the RCC, terrestrial biome is expected to have more influence in smaller streams.

Ecoregion Scale

We expected the same ecoregion in different countries to be more similar than different ecoregions within each country. We were surprised to find that the only differences among metabolism rates by ecoregion were in relation to the US Terminal Basin which was physically more similar to a montane steppe in lower order sites before transitioning to more xeric characteristics. Vegetation and climate generally explained most of GPP and ER. The substrate and cover metrics in the Grassland Steppe models likely represent shading and flow magnitude.

Climate variables did not directly appear in any best ecoregion-scale models of GPP, but vegetation dominated in all best GPP models at this scale, supporting applicability of biome frameworks (Dodds et al. 2015; Dodds et al. 2019). We only looked at three ecoregion types within the temperate steppe biome. These ecoregion types are probably more closely related than many others (e.g. tropical, arctic), and a broader analysis could further separate out factors controlling metabolism.

Contrasting Continents

Mongolia and the United States vary in development types, geologic timelines, and climates. Climate was most explanatory for our broadest model, likely due to the dramatic separation between Mongolia and the United States by temperature and precipitation. The interaction between backwater presence (as percentage of reach) and ecoregion explained nearly half of GPP, which we propose may be due to the ecoregion-specific presence of backwater as indicative of additional light, as well as presenting a stable habitat for establishment of photosynthesizers and increased residence time. Riparian metrics that we would expect to be proxies for light - for example, mean canopy cover of large trees - were more explanatory in some models than more direct light measures, like mean canopy at bank or mid-channel, such as in the MS for ER, suggesting that our light variables were not at a large enough scale to capture upstream processes. The relative importance of canopy versus other riparian metrics has been linked to spiraling length rather than respiration (Reisinger et al. 2019) but may reflect the variation in the data relating to light-constrained GPP or impacts associated with grazing and other adjacent disturbance. Substrate cover - particularly large woody debris - was uneven by country, as large woody debris presence was infrequent in Mongolia due to limited tree growth and use as fuel where available. These results are consistent with previous work by Hosen et al. (2019) which found similar interactions among climate and proxies for light and disturbance as explanatory of their metabolism estimates.

We were surprised to see that human influence variables across disciplines were not explanatory of US GPP, and attribute this to human influence more evenly applied across US sites. This is similar to our interpretation of climate not appearing as explanatory in best models of Mongolian GPP or ER due to the relatively consistent climate across Mongolia, and to some

degree biased by our method of picking study reach locations to minimize adjacency to direct human influence. Using a term for "country" is likely a proxy for climate and land use. We had hypothesized that human variables and hydrogeomorphology would be most important. They were important in some models, but never explained the majority of variation in GPP or ER. The inconsistency in the best correlated variables within a country's ecoregions - for example, only one variable present as a top variable for a rate in both countries (bankfull height, terminal basin) - indicates that these are functionally different systems, even if rates of metabolism are not significantly different between them.

The recent urbanization and increased grazing in Mongolia has changed patterns of impact to steppe rivers. Rates of metabolism in Mongolian rivers and other rates of metabolism, and most of our variables, were much less variable, enabling more confident extrapolation of results to other Mongolian systems. Rates of metabolism in the US and Mongolia had opposite relationships with some variables, particularly metrics of development; for example, mean canopy cover was significantly positively related to GPP and ER (expressed as a positive value) in the US but had a nonsignificant negative correlation with it in Mongolia, potentially related to a general lack of riparian corridors as a result of grazing up to the bank.

Conclusion

This paper provides the largest metabolism dataset for the metabolism of Mongolian rivers that we are aware of, alongside that of matched ecoregions in the United States. We evaluated patterns and explanatory variables using hydrogeomorphic and land use data on reach-to-watershed, 10-minute to 40-year scales, and identified how we can best predict metabolism across this dataset with scale-specific approaches. These data help fill a gap in the understanding of aquatic systems in grasslands, the temperate steppe, and Central Asia. Filling this gap is

important because Mongolian rivers are particularly vulnerable due to inland high elevation, low precipitation, and snowmelt driven hydrology in a rapidly warming area. Much research has been done in the temperate West, particularly in the United States and Europe, but we examined how application of expectations from this region do not hold up in Mongolia, and provided better context for data worldwide.

We sought to understand what drives differences in metabolism at different scales. Our data suggest that synoptic studies of river metabolism should be designed to sample sites with a range of widths and orders, and consider a wider data context including vegetation, human impact, and substrate characteristics. At an ecoregion level climate was variably but indirectly (vegetation) if not directly explanatory. At the country scale most explanatory variables were inconsistently related to metabolic rates, but hydrogeomorphology was always present in our models. Thus, sampling designs must match the scale of interest: at the local scale, where you sample matters, and at the ecoregional scale, local climate and vegetation need to be considered in sampling design that aims at capturing ecoregional variations. There is a complicated relationship between climate, development and metabolism, made clearer in the more heterogeneous United States. Globally, climate change influences and forms a feedback cycle with river metabolism and organic carbon dynamics.

Metabolism across all our sites was best explained by 40-year climate data, likely related to the fundamental climatic differences between the western United States and Mongolia. This relationship broke down quickly on finer scales, and vegetation was most important at the ecoregion scale to both rates. Explanatory variables by country and by watershed size scale were also inconsistent, further supporting the applicability of biome frameworks, though hydrogeomorphic variables were consistently present but less important. In developing

metabolism models and study, capturing drivers will require consideration of their variability at relevant scales. Our results support broader metric collection but would not require much additional time or resources.

We set out to broaden the context for metabolism by linking to a large, interdisciplinary dataset, expanding our ability to predict and explain rates of GPP and ER, assess the relative roles of traditional ecological frameworks in light of human impact, and improve our ability to anticipate changes associated with a changing climate. We placed our estimates in data contexts familiar to engineers, managers, and macroscale researchers, and identified additional metrics that should be added to the limnologist's toolkit.

Chapter 4 - Warming Affects Biome-Specific River Metabolism of the Mongolian and United States Steppe

Abstract

River metabolism is an essential component of global biogeochemical processes. With more data becoming available, large-scale models of metabolism are possible, extrapolating from numerous local measurements. However, we lack mesoscale models of factors controlling metabolism, including temperature. Therefore, we evaluated the downscaling of broad-scale metabolism models using data collected from broad geographic regions of Mongolia and North America. The understudied rivers of the semi-arid steppe of Mongolia are particularly vulnerable to climate change due to high altitude and latitude. This steppe has matching ecoregions with the United States Great Plains, allowing cross-continent investigation of temperature effects on river metabolism. We evaluated how a broad-scale modeling approach applies at the ecoregion level, projecting changes in estimated rates of metabolism under different warming scenarios. Temperature was not the primary explanatory variable, but directly and indirectly (e.g., in constraining vegetation) influenced modeled rates of metabolism. Our metabolism models did not scale down well when using all data, analyzing by country, and analyzing by ecoregion in specific countries. The Grassland Steppe was the most temperature-sensitive ecoregion for both rates on both continents. Our results suggest that researchers, managers, and policymakers should explicitly consider the scale of interest if they are interested in modeling the influence of increased temperature on river metabolism.

Introduction

Global climate change affects stream ecosystems. Elevated warming from climate change does not lead to consistent changes in temperature across the globe. (Ojima et al. 2019). For example, there are regional differences across altitude and latitude. At a smaller spatial scale, individual and even adjacent watersheds may variously cool or warm depending on their location (Surenkhorloo et al. 2021). While global-scale projections are useful for large-regional scale predictions, these predictions may obscure or be buffered by important local effects. Local ecosystem contingencies could interact with this variable downscaling of climate models to create unique predicted regional ecosystem changes. Lotic systems are no exception to this; the importance of local and upstream processes to stream function (such as respiration and decomposition) interferes with generalizing predictions of global warming and its ecological consequences (Bernhardt et al. 2017; Follstad Shah et al. 2017).

A fundamentally important process within aquatic ecosystems, river metabolism consists of autochthonous production (Gross Primary Production, GPP) and oxidation of allochthonous (generated out of system) and autochthonous (generated within system) carbon (Ecosystem Respiration, ER). Metabolism integrates mechanistic understanding of ecosystem function over the upstream watershed, across types and variables, with carbon as the common ecological currency (Levi & McIntyre 2020; Gounand et al. 2020). Globally, rivers play a major role as biogeochemical reactors acting both as transporters of upstream materials to downstream water bodies (e.g. lakes and oceans) and as a substantial source of greenhouse gasses including carbon dioxide, methane, and nitrous oxide (Li et al. 2021). Different systems (i.e., terrestrial vs aquatic) and scales (including reach to continent as well as daily to 40-year) can vary in drivers and constraints of metabolism, so likely vary in response to warming.

Understanding the controls of river metabolism at the broader biome scale remains a challenge as local-to-watershed level processes lead to fundamental controls on this process. While there is uncertainty regarding the local temperature regulation of aquatic carbon cycles (Jankowski & Schindler 2019), climate can control a range of stream properties (Dodds et al. 2015). Interrelated temperature, PET, and water availability can form a strong positive feedback cycle in arid landscapes (Malsy et al. 2017; Jordan et al. 2018) leading to potential shifts including deforestation, woody encroachment, and desertification (Asner et al. 2004). Given that ecosystem function can depend upon watershed processes, accounting for changes in drivers across biomes demands a broader approach (Gounand et al. 2020).

Global analyses of metabolic responses to temperature have allowed progress in prediction of global warming effects on the components of metabolic processes. Respiration is less sensitive to temperature than predicted by the metabolic theory of ecology (Jankowski & Schindler 2019), as carbon use rates vary with quality, temperature sensitivity, and the ability of organisms to adapt to different temperature regimes (Jankowski & Schindler 2019; Follstad-Shah et al. 2017). Temperature dependence is likely not consistent, linear, or monotonic; Song et al. (2018) found asymmetric convergence, wherein streams with high temperature and GPP/ER are expected to exhibit decreased GPP/ER with warming, and vice versa, not resulting in consistently increased heterotrophy. Rates of metabolism respond differently within different ranges of variables. For example, Bernot et al. (2010) found that ER was more variable, inconsistent, and explained by more complicated models, while GPP was more directly filtered by riparian status. The generality of global analyses can obscure local contingencies related to temperature responses. As such, synthesis of metabolic rate measurements in streams has considered biome-specific effects (Bernhardt et al. 2017). We therefore contrast potential

responses to temperature increases with detailed data on hydrology and natural and anthropogenic characteristics of watersheds across natural climatic gradients (from dry terminal basins, to montane steppe, and grasslands) in each country.

Mongolia is undergoing rapid climatic change (Hofman et al. 2016) with air temperatures increasing twice as quickly as the global average (Dorjsuren et al. 2018) and interacting with rapid land use change (Jordan et al. 2018). The United States is also experiencing dramatic effects of climate change alongside steady intensification of many anthropogenic impacts. In both countries, climate-related drought has exacerbated effects of deforestation and agriculture particularly in the form of grazing, by forming a positive feedback towards desertification between water availability and primary producers' use of it (Silcock and Fensham 2019). These features could complicate application of global models to predict specific responses of river metabolic rates to temperature increases in Mongolia and the United States. Prior analyses of data from the 89 sites across the two countries indicated that factors controlling metabolic rates in rivers varied with scale (watersheds, biomes, countries, see Chapter 3). This paper extends those analyses to explore potential sensitivities of metabolism to temperature increases and contrasts how sensitivities vary with scale.

Materials

Site Descriptions

We evaluated sites of three ecoregions of the Mongolian and United States Great Plains temperate steppe as described by Olson et al. (2001), and selected sites using the RESonate framework to ensure a range of representative sites and conditions based on within-watershed patterns of geomorphology (Thorp et al. 2006). We used diurnal trends of dissolved oxygen and additional variables to model metabolic rates, and to measure auxiliary variables with the field

methods described in Chapters 2 and 3. This included temperature correction to 20º C to account for transient sampling conditions to allow comparison of basal metabolic rate as is common in physiological studies using temperature corrections for GPP and ER using Equation 4.1 based on data

Equation 4.1. Temperature correction of Metabolism rate estimates following Song et al. (2018)

$$rate\ at\ 20ºC = \frac{2.523 * e^{(0.0552*20)}}{2.523 * e^{(0.0552*t_1)} * rate\ at\ t_1} \quad (1)$$

in streams published by Song et al. (2018). We evaluated these rates in the context of additional datasets including 180 hydrogeomorphic metrics collected and calculated as in Kaufmann (1999), in addition to remotely sensed watershed-scale land use, geomorphic, and climate data. Sample variables include substrate characterization, in-channel and riparian area vegetation description, proximity to roads or bridges, 40 year mean temperature and precipitation, percentage of the watershed in cropland, entrenchment ratio, etc. We removed variables which were mathematically redundant (i.e. x and 1-x), after determining which had the higher Spearman's ρ correlation with metabolism during variable selection.

Modeling Approach

We used iterative data refinement from stepwise regression and AIC, but forced temperature into each ecoregion-scale model. We built a mixed model with the particular river sampled as a random effect to account for autocorrelation among multiple sites on a given river. We identified the top ten additional terms to evaluate as ranked by Spearman's ρ correlations with GPP or ER, kept the three top terms (including 40 year mean temperature) based on the highest R^2 through stepwise regression, and then selected the final model based on the lowest value from a stepwise AIC. For the Mongolian grassland, we used a linear model as there was only one river sampled, so river could not be used as a random effect. We used the

predictInterval() function (Knowles and Frederick, 2020) to model changes in rates of metabolism (means and confidence interval) for each of these best models at the within-country ecoregion scale based with a projection range of 0, 1, 1.5, and 2 °C according to the range of most common predictions by 2050 (IPCC, 2019). We used the predict() function instead of the predictInterval() function for the Mongolian Grassland sites, represented by a linear model as they were all within one level of the random effect in the other models (R Core Team, 2020). We build on the work of Song et al. (2018) using a mixed model to predict metabolism based on global data, but refine our modeling to the ecoregion scale and add the additional explanatory variables. Specific values are available in Supplementary Table 4.1. We present respiration as positive values, but figures are trimmed to possible ranges (values above zero, $0 < GPP/ER \leq 3$) despite some projections ending up with negative estimates which indicated respiration was creating dissolved oxygen, which is not possible.

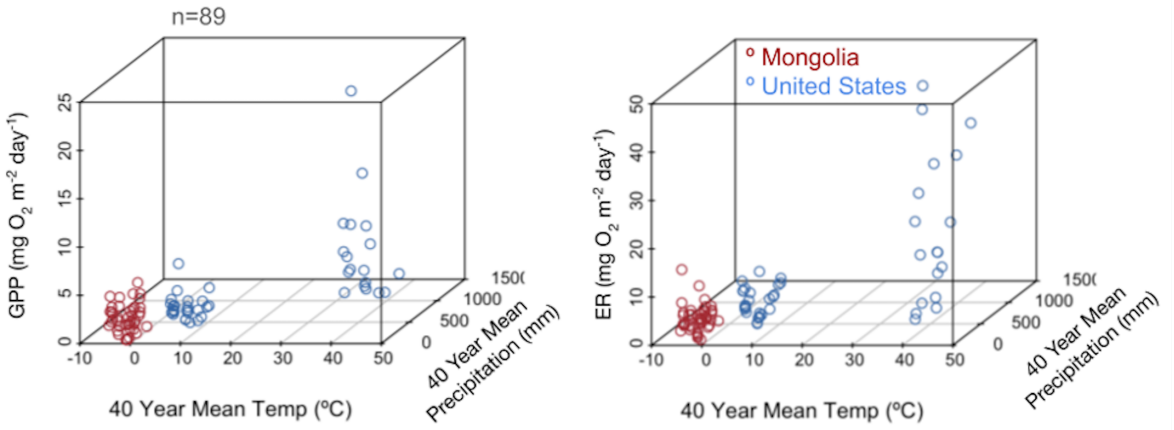


Figure 4.1. GPP and ER by 40 Year Mean Temperature and Precipitation at the watershed scale. Sites in Mongolia are shown in red, while sites in the United States are in blue.

Results

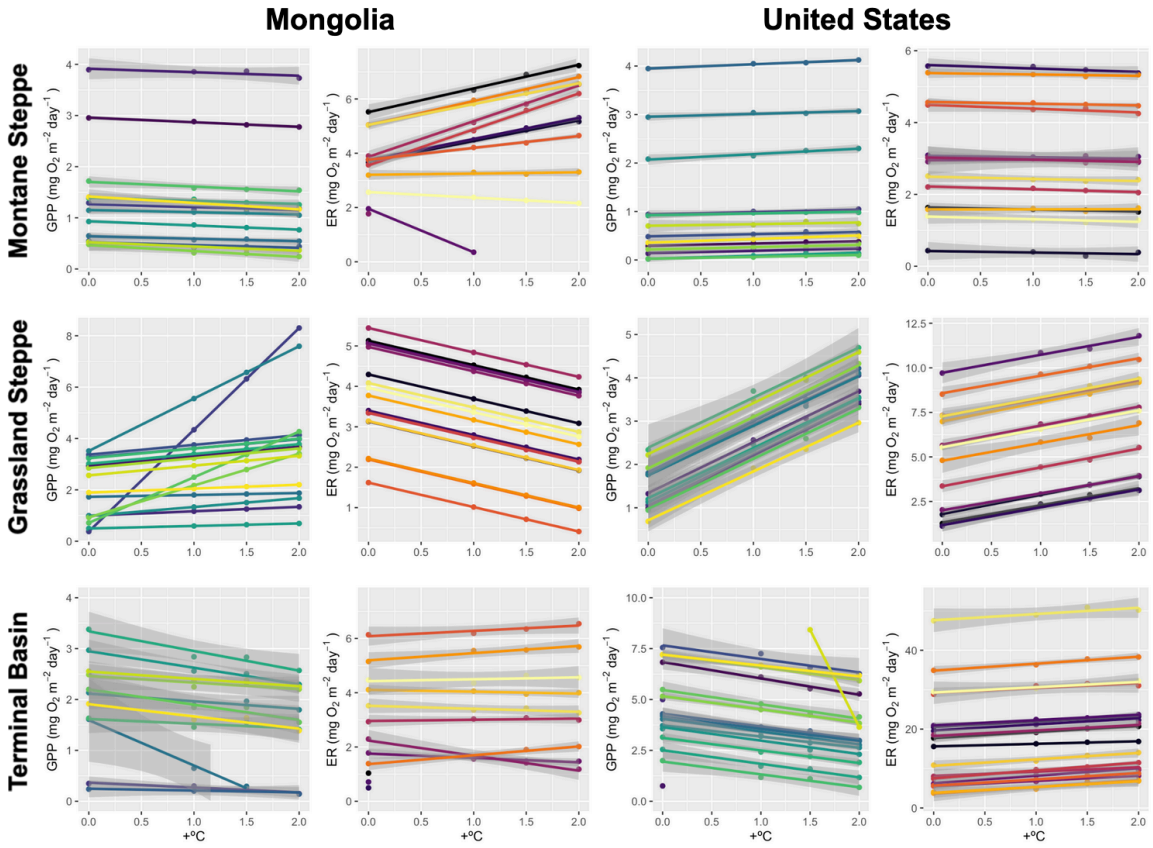


Figure 4.2. Change in individual site rate estimate by subset over up to +2°C using predictInterval() or predict() in conjunction with the best ecoregion-scale model. Scales vary. Note ER is represented as positive, and plots are trimmed to possible values (nonnegative values where ER is represented as positive, 0 < GPP/ER ≤ 3). Each line and greyed confidence interval is a linear smooth per individual site.

Rates for GPP and ER over 0, +1, +1.5, and +2ºC scenarios for each ecoregion in the US and Mongolia indicate that responses to temperature increases vary by metabolic response (GPP or ER), country, and ecoregion (Fig. 4.2). Many of the predicted rates increased with temperature as would be expected with standard physiological responses with temperature. However, some of our models predicted decreases in rates with temperature increases. Generally, responses within an ecoregion and a country were similar, though in a few cases predicted responses deviated from those of other relatively nearby rivers. The decreases with increased temperature were generally related to interactions with other variables. Only one site had

significantly different estimates between the 0, 1ºC scenarios due to the large relative error associated with estimates. There were variably dramatic percentage changes by temperature scenario, though with large uncertainty (Figure 4.2).

Modeled Mongolian Terminal Basin GPP decreased with temperature while ER was variable, but both rates decreased with temperature in the United States (median GPP decreased 1.12 mg O_2 m^{-2} day^{-1} or 26%, ER decreased a comparable bulk quantity 1.06 mg O_2 m^{-2} day^{-1} but only 6%). Montane Steppe GPP had opposite responses in the US and Mongolia - increasing in the US while decreasing in Mongolia - while ER decreased with temperature in the United States and responded variably in Mongolia. Grassland steppe site models predicted clear increases in both rates with temperature, while a less clear increase in GPP, and a clear decrease in respiration. The highest median percentage change- temperature sensitivity- for both rates and countries were in the Grassland Steppe sites. The absolute value of median change ranged from 2% to 19% for ER and 7% to 47% for GPP, though the medians of these changes were similar for both rates (11.36 and 11.58 for ER and GPP respectively). GPP models had slightly higher R^2 values than ER (0.80 vs 0.71 respectively), but GPP was generally an explanatory term in ER models. Specific model information and additional terms are detailed in Table 4.1.

The grassland steppe was the most temperature sensitive ecoregion (highest percentage change, Table 4.2) for both countries and rates, while the Montane steppe was the least. Even in the grassland steppe, temperature was never the most explanatory variable. Variation increased where the model included interactions, with no converging models.

Temperature interacted with vegetation (including large woody debris, non woody understory, and bank canopy) and hydrogeomorphology (wetted width, incision height, and slope). In the Mongolian terminal basin, variation in GPP was explained by temperature only in

the form of an interaction with wetted width, which provided nearly half the model's explanatory power (47%) but had a slope near zero. Mongolian Grassland Steppe GPP was explained by temperature, and by its interactions with non-woody understory and bank canopy. The sign of the relationship between temperature and rate was inconsistent, even within a model: Mongolian Terminal Basin ER was negatively related to temperature but positively and on the same order of magnitude as the interaction between temperature and incision height.

Net ecosystem production, as represented by GPP/ER (Fig. 4.3) also changed variably in direction and magnitude, even within an ecoregion. Specific values are available in Supplementary Table 4.1. Most clearly, the relative autotrophy was predicted to increase in the US Grassland steppe, and shift towards heterotrophy in the MN Montane Steppe and US Terminal Basin, but patterns were not clearly changing in either direction in the US Montane Steppe and MN Terminal Basin, where various sites were predicted to respond differently.

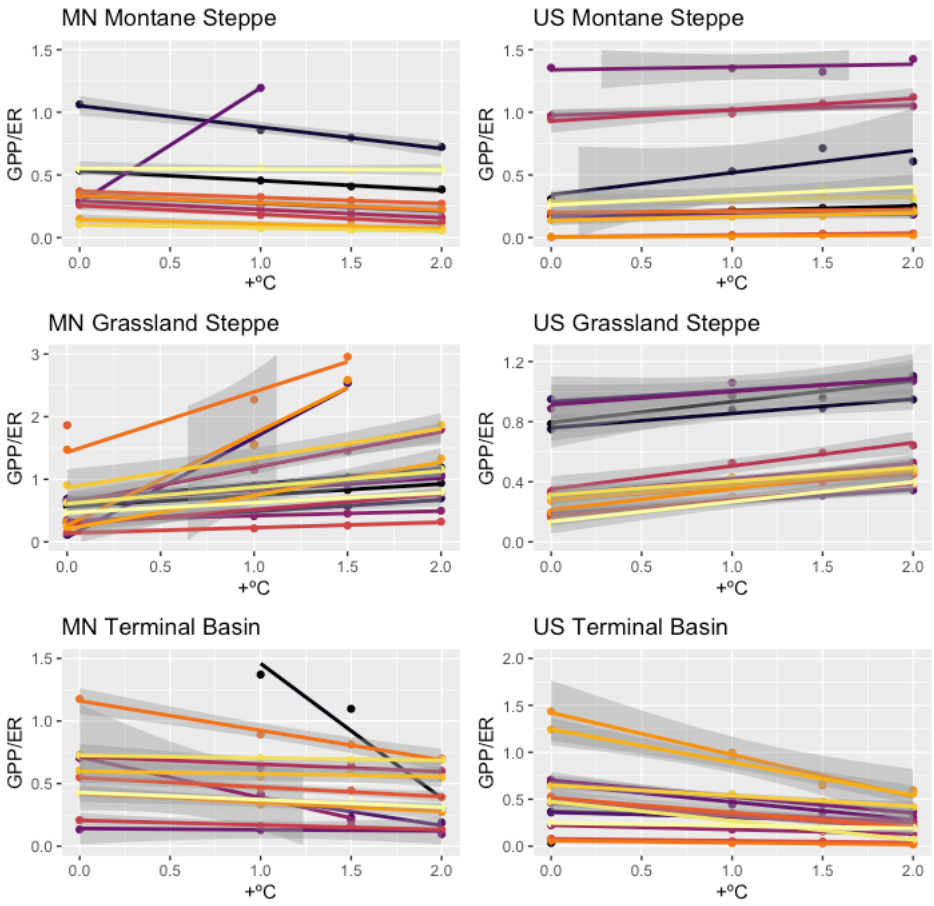


Figure 4.3. GPP/ER (Corrected to 20ºC) by 40 Year Mean Temperature over +0-2ºC Scenarios by Ecoregion using predictInterval() or predict() in conjunction with the best ecoregion-scale model. Plots are trimmed to possible values (nonnegative values where ER is represented as positive, $0 < GPP/ER \leq 3$). Each line and greyed confidence interval is a linear smooth per individual site.

Discussion

Our analyses suggest that context-dependent temperature responses will in part modulate river ecosystem responses to global change. However, we note that over long time periods, other factors in addition to temperature may have primacy. Temperature was not the most explanatory variable, but was important nonetheless. Contrary to expectations, modeled influence of temperature did not categorically increase metabolic rates. Temperature was never a significant term on its own, and often had a small slope (between 0 and 1) when entered into

complex multivariable models. Large and long scale temperature can affect metabolism estimates, but the influence can be indirect and context dependent despite temperature correction of rates to the same water temperature. Our models included predicted negative relationships with temperature, which have been found previously in watersheds in Mongolia and elsewhere (Chen et al. 2018). The variables temperature interacted with included in channel LWD, wetted width, incision height, slope, non-woody understory, and bank canopy - all variables suggesting that light modulates the relationship between temperature and rates of metabolism. Findings were consistent with the United States serving as a temperature projection for Mongolia and as climate characteristics shift towards other biomes with global change.

Semi-arid landscapes are extremely vulnerable to changing climate and land cover, and Mongolia is particularly vulnerable given it is inland and high altitude (Jordan et al. 2018, Dorjsuren et al. 2018, Malsy et al. 2017). Increased temperature and precipitation alter recharge regimes (Dechinlkhundev et al. 2021), while regional climate (with parent geology) control hydrology (Dodds et al. 2019; Poff et al. 1997). The dominance of hydrologic variables across our models outweighs the variability in temperature at the ecoregion level, but is constrained by temperature-related variables including changing precipitation and evapotranspiration patterns. Drought can variably increase (Bernhardt et al. 2017) or decrease (Gounand et al. 2020) GPP and therefore net production. Further, glaciers in Mongolia provide the bulk of freshwater resources but are rapidly shrinking, leading to predicted dramatic shifts in the hydrologic regime (Orkhonselenge and Harbor, 2018) which will influence sources, timing, and transport paths of organic matter (Jankowski & Schindler 2019). The urgency of these pressures is modulated by geomorphic characteristics, additionally constraining the temperature sensitivity of metabolism (Jankowski & Schindler 2019).

Drought could have very strong influences on stream ecosystems in Mongolia because drying is exacerbated by high evapotranspiration within country, with about 90% of precipitation evaporated and not available to subsidize runoff or infiltration and expected to increase (Hoffman et al. 2015). Mongolia has the highest portion of grazed land of any country (80%, Asner et al. 2004). However, rising temperatures are making once-marginal land only suited for grazing more feasible for farming (Jordan et al. 2018; Asner et al. 2004). Water withdrawal to support agriculture and mining as well as urbanization will also accelerate drying (Jordan et al. 2018). Drought is the eminent policy concern in Mongolia, where rangeland is already most often estimated at 70% degraded (Jamsranjav et al. 2018). Though the relative grazing pressure in Mongolia has been increasing since release from Soviet controls on herd sizes, the landscape has been shaped by managed grazing for thousands of years. Mongolia's warming has been more rapid than in the United States Great Plains, where intensive management of livestock has only altered regions in the past several hundred years as compared with at least since 1300 BC (Chen et al. 2018), but rapid development has escalated pressures related to grazing as well as other anthropogenic impacts.

The Grassland Steppe was the most sensitive to temperature, while the Montane Steppe was least likely to change under different temperature scenarios using these models. We were surprised that the Terminal Basin ecoregion, with the greatest range of temperature and metabolism values, and the most difficulty in modeling due to high aeration, had an intermediate temperature relationship and response. This may be due to the range of responses and conditions, but may just reflect the dependence on other explanatory controls.

Predictions on timeframes of these models can't internalize the shifting of these biome boundaries that we expect on much longer timeframes from parallel changes in climate and

anthropogenic impact. Rates of metabolism - and changes in them - vary by country, so we expect that changes in biome will result in changes in carbon dynamics and flux of surface waters. Particularly, as we saw different ecoregions have variable temperature sensitivity, as a particular area shifts towards warmer or wetter, we can expect its temperature sensitivity to shift as well.

Conclusion

Mixed models based on the variation at the ecoregion scale did not support temperature directly affecting rates of river metabolism within steppe ecoregions of Mongolia or the United States, but interactions with temperature did. As climate shifts, however, ecoregions - with distinct metabolic regimes - will as well. Temperature does affect the dominant variables present in these models - particularly hydrologic, vegetation, and land use variables, and can be expected to change the controlling variables we found to be most important. Our modeling results varied by scale, and our metabolism models did not scale down well, but finer scale models are most relevant to management issues and allow identification of areas of concern and evaluations of locally dominant and explanatory variables.

Temperature didn't directly explain the bulk of metabolism rate estimates, but it is important directly and indirectly, constraining many of the primary explanatory variables. Development and resource management decision making alike must consider local variability under a broader general framework - both spatiotemporally and with regard to disciplinary approach. Local use and variability interfere with predictability of management response. Much research on stream ecosystem processes, water quality, ecosystem services, and conservation is based on highly studied temperate zone rivers. We considered understudied Mongonian rivers, and found substantial divergence in results from those obtained in corresponding ecoregions in

North America. This suggests that downscaling of predictions of influence of climate change on streams will require a more nuanced understanding of the contingencies influencing river metabolism in various biomes world-wide.

Acknowledgements

We wish to thank Khaliun Sanchir, Lane Lundeen, Molly Fisher, Samantha Grieger, Andrew Newcombe, Jesús Gomez, Bonnie Teglas, Kari Snelding, Emily Ulrich, and Taylor Hall for field contributions, and Sudeep Chandra for contributing during the editing process. This research was supported by National Science Foundation grants (NSF Macrosystems 1442562 and 1442544 and NSF EPSCOR OIA-1656006). We are aware of no conflicts of interest.

Table 4.1. Summary of each mixed model. LWD = Large Woody Debris

Ecoregion	Rate	Term	Est.	p-value	% Explained
US TB	GPP	(Intercept)	29.92	0.39	
		40 Year Mean Temperature	-0.64	0.48	8.06
		% Bare Ground	-56.80	0.12	32.12
		In-Channel LWD	32454.53	0.01	34.46
		40 Year Mean Temp. : In-Channel LWD	-879.13	0.01	11.25
		% Bare Ground : LWD	-13062.59	0.02	14.11
US TB	ER	(Intercept)	-51.51	0.43	
		40 Year Mean Temperature	1.46	0.41	21.37
		Cobble	124.84	0.00	46.88
		WetWidth	0.06	0.93	18.36
		Cobble:WetWidth	-6.13	0.04	13.39
MN TB	GPP	(Intercept)	-6.77	0.21	
		40 Year Mean Precipitation	0.06	0.07	53.32
		40 Year Mean Temp. : Wetted Width	0.00	0.11	46.68
MN TB	ER	(Intercept)	5.09	0.80	

		40 Year Mean Temperature	-2.59	0.17	9.94
		Incision Height	-5.98	0.63	66.58
		Bankfull Height	-14.23	0.39	13.73
		Incision Height : Bankfull Height	20.41	0.22	6.21
		40 Year Mean Temp. : Incision Height	3.03	0.23	3.53
US MS	GPP	(Intercept)	-3.54	0.06	
		Bank Canopy	0.05	0.04	42.82
		Riparian Ground Cover	7.66	0.01	43.60
		40 Year Mean Temperature	0.06	0.49	1.88
		Bank Canopy : Riparian Ground Cover	-0.12	0.06	11.70
US MS	ER	(Intercept)	2.52	0.04	
		LWD Volume	-71.30	0.14	59.58
		Undercut Bank	2.81	0.05	40.04
		40 Year Mean Temperature	-0.05	0.77	0.38
MN MS	GPP	(Intercept)	0.66	0.95	
		Longitude	-0.01	0.31	82.17
		Bankfull : Depth	0.03	0.16	17.43
		40 Year Mean Temperature	-0.08	0.99	0.39

MN MS	ER	(Intercept)	25.92	0.06	
		Slope	-1.00	0.12	62.33
		Road Proximity	8.49	0.11	13.57
		Temp	4.36	0.13	8.90
		Slope : 40 Year Mean Temp.	-0.19	0.16	15.20
US GS	GPP	(Intercept)	-7.34	0.53	
		Wetted Width	0.00	0.40	44.65
		Cultivated Watershed %	-4.74	0.10	28.17
		40 Year Mean Temperature	1.17	0.42	27.18
US GS	ER	(Intercept)	-2.65	1.00	
		Natural Cover or Shading	2.25	0.01	81.18
		Watershed Slope	-0.74	0.77	14.41
		40 Year Mean Temperature	1.01	1.00	4.40
MN GS	GPP	(Intercept)	3.68	0.00	17.04
		Non-Woody Understory	0.67	0.89	17.04
		Bank Canopy	-0.18	0.01	4.19
		40 Year Mean Temperature	0.38	0.47	11.20
		Non-Woody Understory : Bank Canopy	0.87	0.05	8.02

		Non-Woody Understory : 40 Year Temp.	6.78	0.10	5.26
		Bank Canopy : 40 Year Mean Temp.	-0.03	0.26	37.25
MN GS	ER	(Intercept)	0.26	0.83	36.03
		Velocity	5.44	0.01	10.32
		Bare Ground	172.82	0.04	9.39
		40 Year Mean Temperature	-0.60	0.12	44.25

Table 4.2. Summary of rate estimates by scenario and model

	Median C			
	0.00	1.00	1.50	2.00
MN GS ER	3.78	3.17	2.87	2.57
US GS ER	5.54	6.56	7.15	7.58
MN MS ER	3.71	4.49	4.89	5.24
US MS ER	3.08	2.94	2.90	2.93
MN TB ER	3.23	3.20	3.25	3.13
US TB ER	16.66	17.72	18.37	18.77
median ER	3.75	3.85	4.07	4.19
MN GS GPP	2.56	3.23	3.42	3.67
US GS GPP	1.32	2.57	3.06	3.69
MN MS GPP	1.22	1.13	1.14	1.11
US MS GPP	0.48	0.53	0.58	0.56
MN TB GPP	1.77	1.53	1.59	1.40
US TB GPP	4.64	3.42	3.24	2.86
median GPP	1.55	2.05	2.33	2.13

Chapter 5 - Conclusion

River metabolism is increasingly accessible from analyses of monitoring datasets, and is used as a multiscale functional indicator integrating carbon transformations forming the basal food web as well as over the upstream drained area. I was interested in improving the quality and comparability of metabolism estimates and evaluating sources of bias in Chapter 2. I expanded the extent of available metabolism estimates in Chapter 3 to three temperate steppe ecoregions of Mongolia, and provided data from matched ecoregions in the United States for context and comparison, evaluating how explanatory controls differ by scale. In Chapter 4, I then used the modeling approach from Chapter 3 to project how we would expect rates of GPP and ER to change under a warming climate, and examined the role of temperature in predicting reach-scale metabolism with scale-specific models. Here I summarize the major findings of each chapter, as well as shortfalls and future research directions that would build on this work and bring us towards more accurate and well-understood metabolism research.

Methods are Underreported, but Significantly Change Metabolism Estimates

Our literature review as part of Chapter 2 found that methods and QA/QC are underreported in the literature, and that decision points including equipment type, data source, sensor deployment location, and locally versus remotely sensed data source could systematically bias estimates of metabolism. The largest differences in resultant rate estimates in this assessment were associated with deployment location and accounting for sensor accuracy and drift. ER was more sensitive than GPP to most methodological choices. Better reporting of methods would increase the value and utility of data. Our recommendations are generally simple enough to adopt, and I empirically show that following general operational guidelines and reporting can improve the value and precise comparison of estimates among studies. This work

indicates a need for more careful equipment, field, data, and reporting methods. Cross-sectional matrix deployments in streams in varying biomes and morphologies would allow for finer-scale evaluation of heterogeneity. An audit of long-term monitoring QA/QC protocols and increased metadata requirements for aggregation efforts and literature alike would improve the comparability and value of metabolism data.

Metabolism Models and Explanatory Variables Differ by Scale and Region

In Chapter 3, I explored metabolism data based on the methodological issues identified in Chapter 2, and evaluated modeled representations of metabolism at multiple scales and using a broad and interdisciplinary dataset to both evaluate the cross-scale application of a common modeling approach and identify how controls and their datatypes varied by scale. At the broadest scale, climate explained the bulk of metabolism estimates, with additional importance of land use, though dominant drivers varied by scale and categorical separation. Narrower and wider sites (as a proxy for watershed position) were functionally different, reinforcing the applicability of the River Continuum Concept. Variables that I hypothesized to be most important, like slope or canopy, were not consistently explanatory. The variables that I found to be important were reasonably easy to measure but not generally collected by researchers assessing stream metabolism, underlining a need for a broader, more interdisciplinary approach. This chapter can also serve to help future investigators identify which variables to collect. The dominant data type varied clearly by country; Mongolian stream metabolism was primarily explained by vegetation, while the values from United States sites were most separated by land use. I found no consistent basis for prediction or remote sensing capability of factors that could indicate metabolism directly, though many important variables identified in our models were conducive to remote sensing. Models representing systems are often published as standalone explanatory structures,

but they must be evaluated for value as predictive tools or in broader data contexts. Further work would benefit from placing metabolism estimates in even more comprehensive data (including temporal) contexts, and from evaluating different forms and scales of potential drivers to improve our predictive ability and movement towards accuracy in remote sensing. I only looked at three ecoregions within the temperate steppe, but there are still regions where available data is lacking that should be further studied and evaluated.

Warming Directly and Indirectly Affects Biome-Specific River Metabolism

Having found that broad-scale models poorly explained finer-scale metabolism estimates, I built models in Chapter 3 on the ecoregion scale to evaluate the sensitivity to warming temperatures. Classical biome theory posits that two replicates of the same biome are more functionally similar on different continents than would be two adjacent biomes of different types (the biome concept is based on functional similarities of climate and terrestrial vegetation type regardless of phylogeny). I found that ecoregions were not comparable across continents with respect to factors predicted to influence river metabolism, and that some of this difference is due to human influence overriding predictions simply based on temperature, precipitation, and altitude/latitude expectations, suggesting that similarities among terrestrial biomes do not necessarily extend to river ecosystem properties. Long term shifts in precipitation and temperature related to climate change in a given area will eventually shift the biome as well. For example, a steppe grassland could shift to desert with greater difference between actual and potential evapotranspiration. I did not find support for temperature directly or greatly affecting river metabolism within the MN and US temperate steppe, but it did affect the dominant variables present in these models - particularly hydrologic, vegetation, and land use variables, and can be expected to change the most important variables. This highlights a need for local

tailoring of management efforts to improve outcomes. In particular, application of management plans proven successful in other systems should be used elsewhere with caution, particularly in the most-temperature-sensitive grassland systems. Local use and variability interfere with predictability of response to management. Extending this approach to other biomes would evaluate whether this variability is consistent, and would help identify both particular regions of climate vulnerability, as well as other more sensitive controls on carbon dynamics.

The greatest limitation of this project is that we only had at most a few days at each site, and we lacked the ability to observe longer term trends and variation. Further work will combine these spatial, conceptual, and temporal breadths to understand metabolism better. Modeling capability is limited with respect to overcoming the limitations of a short time deployment and modeling two-station metabolism in a Bayesian framework. An additional limitation to this study is that working in remote areas precluded collection and preservation of nutrient samples and gas samples that others have been used to inform broad scale patterns of river metabolism. Hopefully future technology will enable such collection, allowing scientists to approach questions we could not, including the drivers, patterns, and relative proportion of autotrophic and heterotrophic respiration.

I was fortunate to have the opportunity to build an incredibly broad dataset, in terms of geographic extent as well as in the many disciplines and datatypes covered by the teams on the MACRO project, and one of the unique aspects of this data is having all of those datasets collected at the same place and time and using the same methods and equipment. I am excited for the future use of the growing metabolism knowledge base, and am grateful for the opportunity to have contributed to understanding on river metabolism, increasingly a concept we can use to study and track our waters in a changing world.

References

Appling, A. P., Read, J. S., Winslow, L. A., Arroita, M., Bernhardt, E. S., Griffiths, N. A., … Yackulic, C. B. (2018). The metabolic regimes of 356 rivers in the United States. Scientific Data, 5, 180292. https://doi.org/10.1038/sdata.2018.292

Hebbali, A. (2020). olsrr: Tools for Building OLS Regression Models. R package version 0.5.3. https://CRAN.R-project.org/package=olsrr

Arroita, M., Elosegi, A., & Hall, R. O. (2019). Twenty years of daily metabolism show riverine recovery following sewage abatement. Limnology and Oceanography, 64, S77–S92. https://doi.org/10.1002/lno.11053

Asner, G. P., Elmore, A. J., Olander, L. P., Martin, R. E., & Harris, A. T. (2004). Grazing systems, ecosystem responses, and global change. Annu. Rev. Environ. Resour., 29, 261-299.

BENDA, L., POFF, N. L., MILLER, D., DUNNE, T., REEVES, G., PESS, G., & POLLOCK, M. (2004). The Network Dynamics Hypothesis: How Channel Networks Structure Riverine Habitats. BioScience, 54(5), 413. https://doi.org/10.1641/0006-3568(2004)054[0413:TNDHHC]2.0.CO;2

Bernhardt, E. S., Blaszczak, J. R., Ficken, C. D., Fork, M. L., Kaiser, K. E., & Seybold, E. C. (2017). Control Points in Ecosystems: Moving Beyond the Hot Spot Hot Moment Concept. Ecosystems. https://doi.org/10.1007/s10021-016-0103-y

Bernhardt, E. S., Heffernan, J. B., Grimm, N. B., Stanley, E. H., Harvey, J. W., Arroita, M., … Yackulic, C. B. (2018). The metabolic regimes of flowing waters. Limnology and Oceanography, 63(October), S99–S118. https://doi.org/10.1002/lno.10726

Bernot, M. J., Sobota, D. J., Jr, R. O. H., Mulholland, P. J., Dodds, W. K., Webster, J. R., … Wilson, K. C. (2010). Inter-regional comparison of land-use effects on stream metabolism. Freshwater Biology, 55, 1874–1890. https://doi.org/10.1111/j.1365-2427.2010.02

Bortolotti, L. E., St. Louis, V. L., & Vinebrooke, R. D. (2019). Drivers of ecosystem metabolism in restored and natural prairie wetlands. Canadian Journal of Fisheries and Aquatic Sciences, 76(12), 2396–2407. https://doi.org/10.1139/cjfas-2018-0419

Brett, M. T., Bunn, S. E., Chandra, S., Galloway, A. W. E., Guo, F., Kainz, M. J., … Wehr, J. D. (2017). How important are terrestrial organic carbon inputs for secondary production in freshwater ecosystems? Freshwater Biology, (January), 833–853. https://doi.org/10.1111/fwb.12909

Cardinale, B. J., Palmer, M. A., Swan, C. M., Brooks, S., & LeRoy Poff, N. (2002). The influence of substrate heterogeneity on biofilm metabolism in a stream ecosystem.

Ecology, 83(2), 412–422. https://doi.org/10.1890/0012-9658(2002)083[0412:TIOSHO]2.0.CO;2

Chen, J., John, R., Sun, G., Fan, P., Henebry, G. M., Fernández-Giménez, M. E., … Qi, J. (2018). Prospects for the sustainability of social-ecological systems (SES) on the Mongolian plateau: Five critical issues. Environmental Research Letters, 13(12). https://doi.org/10.1088/1748-9326/aaf27b

Chowanski, K., Kunza, L., Hoffman, G., Genzoli, L., & Stickney, E. (2020). River management alters ecosystem metabolism in a large oligotrophic river. Freshwater Science, 39(February), 534–548. https://doi.org/10.1086/710082

Intergovernmental Panel on Climate Change. (2019). Climate change and land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems. Report, Intergovernmental Panel on Climate Change, Switzerland.

Cox, R. W. (1979). Ideologies and the new international economic order: reflections on some recent literature. International Organization, 33(2), 257-302.

Dechinlkhundev, D., Zorigt, M., & Dorjsuren, I. (2021). The sustainable use of groundwater resources concerning further climate change scenarios in Ulaanbaatar City Area, Mongolia. Journal of Water and Climate Change.

Demars, B. O. L. (2019). Hydrological pulses and burning of dissolved organic carbon by stream respiration. Limnology and Oceanography, 64(1), 406–421. https://doi.org/10.1002/lno.11048

Demars, B. O. L., Thompson, J., & Manson, J. R. (2015). Stream metabolism and the open diel oxygen method: Principles, practice, and perspectives. Limnology and Oceanography: Methods, 13(7), 356–374. https://doi.org/10.1002/lom3.10030

Demars, B. O. L., Russell Manson, J., Ólafsson, J. S., Gíslason, G. M., Gudmundsdóttir, R., Woodward, G., … Friberg, N. (2011). Temperature and the metabolic balance of streams. Freshwater Biology, 56(6), 1106–1121. https://doi.org/10.1111/j.1365-2427.2010.02554.x

Dodds, W. K., & Cole, J. J. (2007). Expanding the concept of trophic state in aquatic ecosystems: it's not just the autotrophs. Aquatic Sciences, 69(4), 427-439.

Dodds, W. K., Bruckerhoff, L. A., Batzer, D., Schechner, A., Pennock, C., Renner, E., … Grieger, S. (2019). The freshwater biome gradient framework: Predicting macroscale properties based on latitude, altitude, and precipitation. Ecosphere, 10(7). https://doi.org/10.1002/ecs2.2786

Dodds, W. K., Gido, K., Whiles, M. R., Daniels, M. D., & Grudzinski, B. P. (2015). The stream biome gradient concept: Factors controlling lotic systems across broad biogeographic scales. Freshwater Science, 34(1), 1-19.

Dodds, W. K., Rose, K. C., Fei, S., & Chandra, S. (2021). Macrosystems revisited: challenges and successes in a new subdiscipline of ecology. Frontiers in Ecology and the Environment, 19(1), 4-10.

Dodds, W. K., Veach, A. M., Ruffing, C. M., Larson, D. M., Fischer, J. L., & Costigan, K. H. (2013). Abiotic controls and temporal variability of river metabolism: multiyear analyses of Mississippi and Chattahoochee River data. Freshwater Science, 32(4), 1073–1087. https://doi.org/10.1899/13-018.1

Dodds, W. K., Higgs, S. A., Spangler, M. J., Guinnip, J. P., Scott, J. D., Hedden, S. C., … Evans-White, M. A. (2018). Spatial heterogeneity and controls of ecosystem metabolism in a Great Plains river network. Hydrobiologia, 813(1). https://doi.org/10.1007/s10750-018-3516-0

Dodds, W. K., Gido, K. B., Whiles, M. R., Fritz, K. M., & Matthews, W. J. (2004). Life on the edge: the ecology of Great Plains prairie streams. BioScience, 54(3), 205. https://doi.org/10.1641/0006-3568(2004)054[0205:LOTETE]2.0.CO;2

Dorjsuren, B., Yan, D., Wang, H., Chonokhuu, S., Enkhbold, A., Yiran, X., … Abiyu, A. (2018). Observed trends of climate and river discharge in Mongolia's Selenga sub-basin of the Lake Baikal basins. Water (Switzerland), 10(10), 1–18. https://doi.org/10.3390/w10101436

Edwards, R. W., & Owens, M. (1962). The Effects of Plants on River Conditions IV. The Oxygen Balance of a Chalk Stream. The Journal of Ecology, 50(1), 207. https://doi.org/10.2307/2257205

Fausch, K. D., Torgersen, C. E., Baxter, C. V., & Li, H. W. (2002). Landscapes to Riverscapes : Bridging the Gap between Research and Conservation of Stream Fishes. BioScience, 52(6), 483–498. https://doi.org/10.1641/0006-3568(2002)052

Feijó-Lima, R., E. Zandonà, B.S. da Silva, F. Tromboni, T.P. Moulton, S.A. Thomas (2019). Longitudinal dimensions of land-use impacts in riverine ecosystems. ACTA LIMNOLOGICA BRASILIENSIA 31, e107

Feijó-Lima, R., S. Mcleay, E.F. Silva-Junior, F. Tromboni, T.P. Moulton, E. Zandonà, S.A. Thomas (2018). Quantitatively describing the downstream effects of an abrupt land cover transition: buffering effects of a forest remnant on a stream impacted by cattle grazing. INLAND WATERS 8:3, 294-311

Follstad Shah, J. J., Kominoski, J. S., Ardón, M., Dodds, W. K., Gessner, M. O., Griffiths, N. A., … Zeglin, L. H. (2017). Global synthesis of the temperature sensitivity of leaf litter breakdown in streams and rivers. Global Change Biology, 23(8), 3064–3075. https://doi.org/10.1111/gcb.13609

Frissell, C. A., Liss, W. J., Warren, C. E., & Hurley, M. D. (1986). A hierarchical framework for stream habitat classification: Viewing streams in a watershed context. Environmental Management, 10(2), 199–214. https://doi.org/10.1007/BF01867358

Gounand, I., Little, C. J., Harvey, E., & Altermatt, F. (2020). Global quantitative synthesis of ecosystem functioning across climatic zones and ecosystem types. Global Ecology and Biogeography, 29(7), 1139–1176. https://doi.org/10.1111/geb.13093

Grace, M., Giling, D. P., Hladyz, S., Caron, V., Thompson, R. M., & Nally, R. Mac. (2015). Fast processing of diel oxygen curves: Estimating stream metabolism with BASE (BAyesian Single-station Estimation). Limnology and Oceanography-Methods, 13, 103–114. https://doi.org/10.1002/lom.10011

Gücker, B., & Pusch, M. T. (2006). Regulation of nutrient uptake in eutrophic lowland streams. Limnology and Oceanography, 51(3), 1443-1453.

Hall Jr, R. O., & Hotchkiss, E. R. (2017). Stream Metabolism. Methods in Stream Ecology: Third Edition (Vol. 2). Elsevier Inc. https://doi.org/10.1016/B978-0-12-813047-6.00012-7

Hall Jr, R. O., & Tank, J. L. (2005). Correcting whole-stream estimates of metabolism for groundwater input. Limnology and Oceanography: Methods, 3(4), 222-229.

Hall, R. O., Tank, J. L., Baker, M. A., Rosi-Marshall, E. J., & Hotchkiss, E. R. (2016). Metabolism, Gas Exchange, and Carbon Spiraling in Rivers. Ecosystems, 19(1), 73–86. https://doi.org/10.1007/s10021-015-9918-1

Hamid, A., Bhat, S. U., & Jehangir, A. (2020). Local determinants influencing stream water quality. Applied Water Science, 10(1), 1–16. https://doi.org/10.1007/s13201-019-1043-4

Heffernan, J. B., Soranno, P. A., Angilletta, M. J., Buckley, L. B., Gruner, D. S., Keitt, T. H., … Weathers, K. C. (2014). Macrosystems ecology: Understanding ecological patterns and processes at continental scales. *Frontiers in Ecology and the Environment*, *12*(1), 5–14. https://doi.org/10.1890/130017

Hoellein, T. J., Bruesewitz, D. A., & Richardson, D. C. (2013). Revisiting Odum (1956): A synthesis of aquatic ecosystem metabolism. Limnology and Oceanography, 58(6), 2089–2100. https://doi.org/10.4319/lo.2013.58.6.2089

Hofmann, J., Tuul, D., & Enkhtuya, B. (2016). Integrated Water Resources Management: Concept, Research and Implementation. https://doi.org/10.1007/978-3-319-25071-7

Hosen, J. D., Aho, K. S., Appling, A. P., Creech, E. C., Fair, J. H., Hall, R. O., … Raymond, P. A. (2019). Enhancement of primary production during drought in a temperate watershed is greater in larger rivers than headwater streams. Limnology and Oceanography, 64(4), 1458–1472. https://doi.org/10.1002/lno.11127

Hotchkiss, E. R., Sadro, S., & Hanson, P. C. (2018). Toward a more integrative perspective on carbon metabolism across lentic and lotic inland waters. Limnology and Oceanography Letters, 3(3), 57–63. https://doi.org/10.1002/lol2.10081

Huryn, A. D., Benstead, J. P., & Parker, S. M. (2017). Seasonal changes in light availability modify the temperature dependence of ecosystem metabolism in an arctic stream Published by : Wiley on behalf of the Ecological Society of America Stable URL : http://www.jstor.org/stable/43493909 Seasonal changes in , 95(10), 2826–2839. https://doi.org/10.1890/13-1963.1

Hynes, H. (1975). The Stream and its valley. Verh. Internat. Verein. Limnol. 19(1), 1-15. https://doi.org/10.1080/03680770.1974.11896033

Irwin, C. E., Culp, J. M., & Yates, A. G. (2020). Spatio-temporal variation of benthic metabolism in a large, regulated river. Canadian Water Resources Journal (Vol. 45). https://doi.org/10.1080/07011784.2020.1724198

Jamsranjav, C., Reid, R. S., Fernández-Giménez, M. E., Tsevlee, A., Yadamsuren, B., & Heiner, M. (2018). Applying a dryland degradation framework for rangelands: the case of Mongolia. Ecological Applications, 0(0), 1–21. https://doi.org/10.1002/eap.1684

Jankowski, K. J., & Schindler, D. E. (2019). Watershed geomorphology modifies the sensitivity of aquatic ecosystem metabolism to temperature. Scientific Reports, 9(1), 1–10. https://doi.org/10.1038/s41598-019-53703-3

Jankowski, K. J., Mejia, F. H., Blaszczak, J. R., & Holtgrieve, G. W. (2021). Aquatic ecosystem metabolism as a tool in environmental management. Wiley Interdisciplinary Reviews: Water, 8(4), e1521.

Jared E. Knowles and Carl Frederick (2020). merTools: Tools for Analyzing Mixed Effect Regression Models. R package version 0.5.2. https://CRAN.R-project.org/package=merTools

Jassby, A. D., & Platt, T. (1976). Mathematical formulation of the relationship between photosynthesis and light for phytoplankton. Limnology and oceanography, 21(4), 540-547.

Jordan, G., Goenster-Jordan, S., Lamparter, G. J., Ulziisuren, B., Soninkishig, N., Schlecht, E., & Buerkert, A. (2018). Water use in agro-pastoral livelihood systems within the Bulgan River watershed of the Altay Mountains, Western Mongolia. Agriculture, Ecosystems and Environment, 251(October 2017), 180–193. https://doi.org/10.1016/j.agee.2017.09.008

Judd, C., Thom, R., Borde, A., Woodruff, D., Roegner, C., Vavrinec, J., ... & Zhang, J. (2009). Eelgrass enhancement and restoration in the lower Columbia River estuary. Pacific Northwest National Laboratory, prepared for the US Department of Energy.

Julian, J. P., Doyle, M. W., & Stanley, E. H. (2008). Empirical modeling of light availability in rivers. Journal of Geophysical Research: Biogeosciences, 113(3), 1–16. https://doi.org/10.1029/2007JG000601

Karthe, D., Abdullaev, I., Boldgiv, B., Borchardt, D., Chalov, S., Jarsjö, J., … Nittrouer, J. A. (2017). Water in Central Asia: an integrated assessment for science-based management. Environmental Earth Sciences, 76(20). https://doi.org/10.1007/s12665-017-6994-x

Kaufmann, P. R., Levine, P., Robison, E. G., Seeliger, C., & Peck, D. V. (1999). Quantifying Physical Habitat in Wadeable Streams. EPA/620/R-99/003. U.S. Environmental Protection Agency, Washington, D.C., (July), 130.

Kida, M., Myangan, O., Oyuntsetseg, B., Khakhinov, V., Kawahigashi, M., & Fujitake, N. (2018). Dissolved organic matter distribution and its association with colloidal aluminum and iron in the Selenga River Basin from Ulaanbaatar to Lake Baikal. Environmental Science and Pollution Research, 25(12), 11948–11957. https://doi.org/10.1007/s11356-018-1462-z

Koenig, L. E., Helton, A. M., Savoy, P., Bertuzzo, E., Heffernan, J. B., Hall, R. O., & Bernhardt, E. S. (2019). Emergent productivity regimes of river networks. Limnology and Oceanography Letters, 4(5), 173–181. https://doi.org/10.1002/lol2.10115

Kovach, R. P., Dunham, J. B., Al-Chokhachy, R., Snyder, C. D., Letcher, B. H., Young, J. A., … Muhlfeld, C. C. (2019). An integrated framework for ecological drought across riverscapes of North America. BioScience, 69(6), 418–431. https://doi.org/10.1093/biosci/biz040

Lazorchak, J., Hill, B., Averill, D., Peck, D., & Klemm, D. (2000). Environmental Monitoring and Assessment Program- Surface Waters : Field Operations and Methods for Measuring the Ecological Condition of Non-wadeable Rivers And Streams, 204. Retrieved from https://www.researchgate.net/publication/238086010_Section_2_Overview_of_Field_Operations

Levi, P. S., & McIntyre, P. B. (2020). Ecosystem responses to channel restoration decline with stream size in urban river networks. Ecological Applications, 30(5), 1–17. https://doi.org/10.1002/eap.2107

Levins, R. (1966). The strategy of model building in population biology. American scientist, 54(4), 421-431.

Li, M., Peng, C., Zhang, K., Xu, L., Wang, J., Yang, Y., … He, N. (2021). Headwater stream ecosystem: an important source of greenhouse gases to the atmosphere. Water Research, 190. https://doi.org/10.1016/j.watres.2020.116738

Long, M. H., Rheuban, J. E., Berg, P., & Zieman, J. C. (2012). A comparison and correction of light intensity loggers to photosynthetically active radiation sensors. Limnology and Oceanography: Methods, 10(JUNE), 416–424. https://doi.org/10.4319/lom.2012.10.416

Maasri, A., Thorp, J. H., Gelhaus, J. K., Tromboni, F., Chandra, S., & Kenner, S. J. (2019). Communities associated with the Functional Process Zone scale: A case study of stream macroinvertebrates in endorheic drainages. Science of the Total Environment, 677, 184-193.

Maasri, A., Pyron, M., Arsenault, E. R., Thorp, J. H., Mendsaikhan, B., Tromboni, F., … Boldgiv, B. (2021). Valley-scale hydrogeomorphology drives river fish assemblage variation in Mongolia. Ecology and Evolution, 11(11), 6527–6535. https://doi.org/10.1002/ece3.7505

Malsy, M., Flörke, M., & Borchardt, D. (2017). What drives the water quality changes in the Selenga Basin: climate change or socio-economic development?. Regional environmental change, 17(7), 1977-1989.

Mccutchan, J. H., Lewis, W. M., & Saunders, J. F. (1998). Uncertainty in the estimation of stream metabolism from open-channel oxygen concentrations. Journal of the North American Benthological Society, 17(2), 155–164. https://doi.org/10.2307/1467959

Mulholland, P. J., Fellows, C. S., Tank, J. L., Grimm, N. B., Webster, J. R., Hamilton, S. K., … Peterson, B. J. (2001). Inter-biome comparison of factors controlling stream metabolism. Freshwater Biology, 46(11), 1503–1517. https://doi.org/10.1046/j.1365-2427.2001.00773.x

O'Donnell, B. M., & Hotchkiss, E. R. (2019). Coupling Concentration- and Process-Discharge Relationships Integrates Water Chemistry and Metabolism in Streams. Water Resources Research, 55(12), 10179–10190. https://doi.org/10.1029/2019WR025025

Odum, H. T. (1956). Primary Production in Flowing Waters. https://doi.org/10.1109/iscas.1994.408763

Ogle, D.H., P. Wheeler, & Dinno, A. (2021). FSA: Fisheries Stock Analysis. R package version 0.8.32, https://github.com/droglenc/FSA.

Ojima, D., Togtokh, C., Galvin, K. A., Hopping, K., Beeton, T., Ulambayar, T., ... & Myagmarsuren, A. (2017). Coping with Climate Extremes in Mongolian Pastoral Communities. In Making Climate Compatible Development Happen (pp. 182-200). Routledge.

Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V. N., Underwood, E. C., … Kassem, K. R. (2001). Terrestrial ecoregions of the world: A new map of life on Earth. BioScience, 51(11), 933–938. https://doi.org/10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2

Orkhonselenge, A., & Harbor, J. M. (2018). Impacts of Modern Glacier Changes on Surface Water Resources in Western and Northern Mongolia. Journal of Water Resource and Protection, 10(06), 559.

Pastor, A., Lupon, A., Gómez-gener, L., Rodríguez-castillo, T., Abril, M., Arce, M. I., … Catalán, N. (2017). Local and regional drivers of headwater streams metabolism : insights from the fi rst AIL collaborative project, 36(1), 67–85. https://doi.org/10.23818/limn.36.06

Pearce, N. J. T., Thomas, K. E., Chambers, P. A., Venkiteswaran, J. J., & Yates, A. G. (2020). Metabolic regimes of three mid-order streams in southern Ontario, Canada exposed to contrasting sources of nutrients. Hydrobiologia, 847(8), 1925–1942. https://doi.org/10.1007/s10750-020-04222-0

Pielke, R. A., & Wilby, R. L. (2012). Regional climate downscaling: What's the point? Eos, 93(5), 52–53. https://doi.org/10.1029/2012EO050008

Poff, N. L., Allan, J. D., Bain, M. B., Karr, J. R., Prestegaard, K. L., Richter, B. D., … Stromberg, J. C. (1997). A paradigm for river conservation and restoration. BioScience, 47(11), 769–784. https://doi.org/10.2307/1313099

Poff, N. L., Allan, J. D., Bain, M. B., Karr, J. R., Prestegaard, K. L., Richter, B. D., … Stromberg, J. C. (1997). The natural flow regime: A paradigm for river conservation and restoration. BioScience, 47(11), 769–784. https://doi.org/10.2307/1313099

Preiner, S., Dai, Y., Pucher, M., Reitsema, R. E., Schoelynck, J., Meire, P., & Hein, T. (2020). Effects of macrophytes on ecosystem metabolism and net nutrient uptake in a groundwater fed lowland river. Science of the Total Environment, 721, 137620. https://doi.org/10.1016/j.scitotenv.2020.137620

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Reichert, P., Uehlinger, U., & Acuña, V. (2009). Estimating stream metabolism from oxygen concentrations: Effect of spatial heterogeneity. Journal of Geophysical Research: Biogeosciences, 114(3), 1–15. https://doi.org/10.1029/2008JG000917

Reiman, J. H., & Xu, Y. J. (2018). Diel variability of pCO 2 and CO 2 outgassing from the lower Mississippi River: Implications for riverine CO 2 outgassing estimation. Water (Switzerland), 11(1), 1–15. https://doi.org/10.3390/w11010043

Reisinger, A. J., Doody, T. R., Groffman, P. M., Kaushal, S. S., & Rosi, E. J. (2019). Seeing the light: urban stream restoration affects stream metabolism and nitrate uptake via changes in canopy cover. Ecological Applications, 29(6), 1172–1186. https://doi.org/10.1002/eap.1941

Riley, A. J., & Dodds, W. K. (2013). Whole-stream metabolism: strategies for measuring and modeling diel trends of dissolved oxygen. Freshwater Science, 32(1), 56–69. https://doi.org/10.1899/12-058.1

Riley, A. J., & Dodds, W. K. (2012). The expansion of woody riparian vegetation, and subsequent stream restoration, influences the metabolism of prairie streams. Freshwater Biology, 57(6), 1138-1150.

Rüegg, J., Conn, C. C., Anderson, E. P., Battin, T. J., Bernhardt, E. S., Boix Canadell, M., … Yackulic, C. B. (2020). Thinking like a consumer: linking aquatic basal metabolism and

consumer dynamics. For Limnology and Oceanography Letters. https://doi.org/10.1002/lol2.10172

Rüegg, J., Chaloner, D. T., Ballantyne, F., Levi, P. S., Song, C., Tank, J. L., … Lamberti, G. A. (2020). Understanding the Relative Roles of Salmon Spawner Enrichment and Disturbance: A High-Frequency, Multi-Habitat Field and Modeling Approach. Frontiers in Ecology and Evolution, 8(February), 1–14. https://doi.org/10.3389/fevo.2020.00019

Sarneel, J. M., Huig, N., Veen, G. F., Rip, W., & Bakker, E. S. (2014). Herbivores enforce sharp boundaries between terrestrial and aquatic ecosystems. Ecosystems, 17(8), 1426-1438.

Siders, A. C., Larson, D. M., Rüegg, J., & Dodds, W. K. (2017). Probing whole-stream metabolism: influence of spatial heterogeneity on rate estimates. Freshwater Biology, 62(4), 711–723. https://doi.org/10.1111/fwb.12896

Silcock, J. L., & Fensham, R. J. (2018). Using evidence of decline and extinction risk to identify priority regions, habitats and threats for plant conservation in Australia. Australian Journal of Botany, 66(7), 541–555. https://doi.org/10.1071/BT18056

Song, C., Dodds, W. K., Rüegg, J., Argerich, A., Baker, C. L., Bowden, W. B., … Ballantyne, F. (2018). Continental-scale decrease in net primary productivity in streams due to climate warming. Nature Geoscience, 11(6), 415–420. https://doi.org/10.1038/s41561-018-0125-5

Surenkhorloo, P., Buyanaa, C., Dolgorjav, S., Bazarsad, C., Zamba, B., Bayarsaikhan, S., & Heiner, M. (2021). Identifying Riparian Areas of Free Flowing Rivers for Legal Protection : Model Region Mongolia, 1–12.

Thorp, J. H., Thoms, M. C., & Delong, M. D. (2008). The riverine ecosystem synthesis. Academic Press.

Tromboni, F., Liu, J., Ziaco, E., Breshears, D. D., Thompson, K. L., Dodds, W. K., ... & Fei, S. (2021). Macrosystems as metacoupled human and natural systems. Frontiers in Ecology and the Environment, 19(1), 20-29.

Vannote, R. L., Minshall, W. G., Cummins, K. W., Sedell, J. R., & Cushing, C. E. (1980). PERSPECTIVES The River Continuum Conce~tl. Canadian Journal of Fisheries and Aquatic Sciences, 37(1), 130–137.

Welti, N., Striebel, M., Ulseth, A. J., Cross, W. F., DeVilbiss, S., Glibert, P. M., … Hillebrand, H. (2017). Bridging food webs, ecosystem metabolism, and biogeochemistry using ecological stoichiometry theory. Frontiers in Microbiology, 8(JUL), 1–14. https://doi.org/10.3389/fmicb.2017.01298

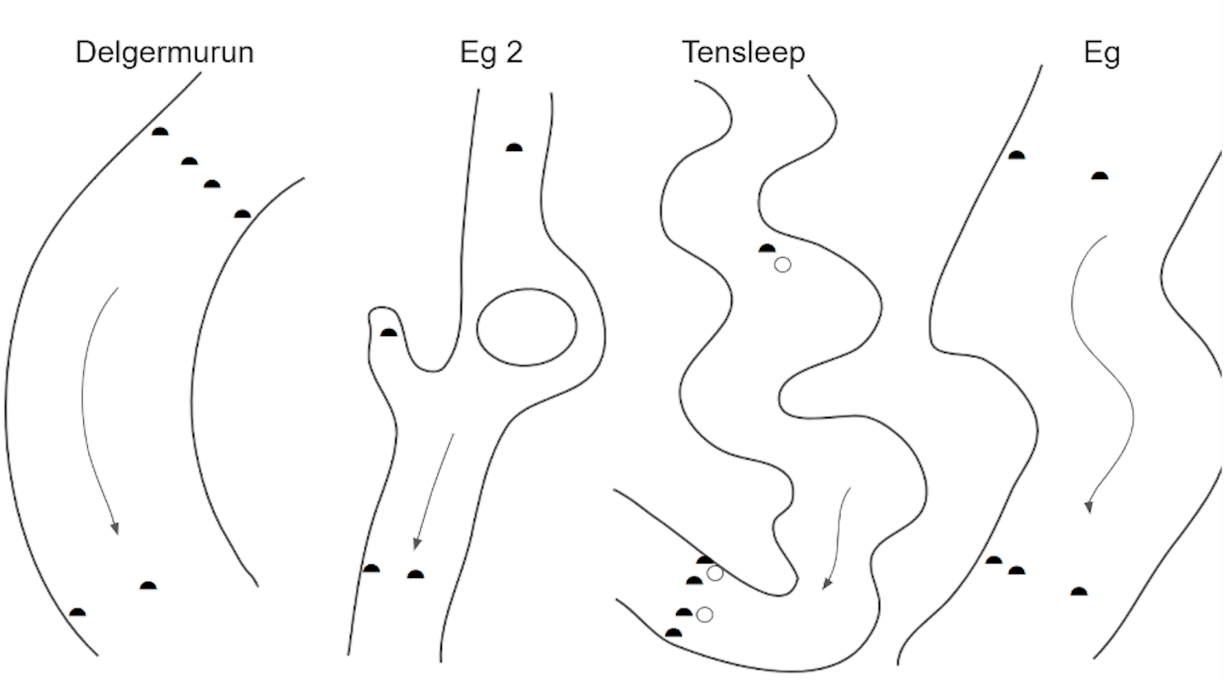
Willett, S. D., McCoy, S. W., & Beeson, H. W. (2018). Transience of the North American High Plains landscape and its impact on surface water. Nature, 561(7724), 528–532. https://doi.org/10.1038/s41586-018-0532-1

Williams, B. S., D'Amico, E., Kastens, J. H., Thorp, J. H., Flotemersch, J. E., & Thoms, M. C. (2013). Automated riverine landscape characterization: GIS-based tools for watershed-scale research, assessment, and management. Environmental monitoring and assessment, 185(9), 7485-7499.

Wolman, M. G., & Miller, J. P. (1960). Magnitude and frequency of forces in geomorphic processes. The Journal of Geology, 68(1), 54-74.

Zimmer, M. A., Kaiser, K. E., Blaszczak, J. R., Zipper, S. C., Hammond, J. C., Fritz, K. M., … Allen, D. C. (2020). Zero or not? Causes and consequences of zero-flow stream gage readings. Wiley Interdisciplinary Reviews: Water, (October 2019), e1436. https://doi.org/10.1002/wat2.1436

Appendix A - Supplemental Information for Chapter 2



Supplemental Figure 2.1. Arrangement of sensors at four reaches described in detail in this manuscript. Open circles indicate a matched deep sensor accompanying a sensor deployed at the surface at the same location. See Table 1 for site and reach characteristics.

Site Description for Detailed Arrays Discussed

Eg

This wide, shallow, rocky river runs next to a town where it serves as a water source and is grazed. The river is more constrained as you move up the reach by unstable, rocky banks, but has no canopy cover. Sensor D is located in a shallow, macrophyte-dominated soft sediment area on the bankside where it might occur to deploy a sensor to avoid losing it in the more turbulent center of the channel. Sensor A is next to sensor D but out of the macrophyte-dominated area with somewhat more coarse substrate. Sensor B is in the thalweg but at the same point along the river length as sensors A and D. Sensor E is in the thalweg further up the reach where more

constrained, while sensor C is at the same location along the reach length but deployed on the side of the channel where more traditionally a sensor would be anchored.

Eg 2

This reach is just above a town and is grazed but has a forested canopy and stable banks with gravel substrate. Sensor A is placed in a backwater under a tree, while sensor B is just downstream on the side of the channel. Sensor C is at the same location along the reach as sensor B, but is deployed in the thalweg. Sensor D is further up the reach and in the thalweg but placed in a deep hole.

Tensleep

This reach is in a meadow draining a glacial lake with alternating shallow riffles and deep pools. There is tree canopy on one side of the river, and incised, grassy banks. Upreach sensors are tied to a fallen, dead tree; G is just under the surface, while H is at the same location but is deployed just off the bottom. Sensors A-F make an array as illustrated in Figure 2.8.

Delgermurun

This is a wide, open, rocky, well grazed and trafficked site. There is no canopy cover, and is reasonably homogeneous in its flow. Sensors A and D are on opposite shorelines where sensors might traditionally be deployed, and F is also on the shoreline, though further upreach. Sensor C is in the thalweg at the same location along the reach as sensors A and D, as is sensor B located between A and C. Sensor E is at the same longitudinal location as F, but is deployed in the center of the channel at depth.

Supplementary Table 2.1. Site locations and pair designations from Figure 2.7. Relative location indicates the placement difference between two sensors in a given pair, Latitude and Longitude are of the reach overall not specific sensors.

Pair Number	Relative Location	Site	Latitude	Longitude
1	center	Delgermurun	49.64	99.92
1	side	Delgermurun	49.64	99.92
2	shallow	Delgermurun	49.64	99.92
2	deep	Delgermurun	49.64	99.92
3	shallow	Eg 2	50.1	101.59
3	deep	Eg 2	50.1	101.59
4	shallow	Kherlen	47	108.84
4	deep	Kherlen	47	108.84
5	shallow	Kherlen 2	46.93	109.42
5	deep	Kherlen 2	46.93	109.42
6	shallow	Kherlen 3	48.05	113.63
6	deep	Kherlen 3	48.05	113.63
7	shallow	Kherlen 4	48.25	108.5
7	deep	Kherlen 4	48.25	108.5
8	shallow	Kherlen 5	48.25	108.5

8	deep	Kherlen 5	48.25	108.5
9	shallow	Kherlen 6	47.73	109.75
9	deep	Kherlen 6	47.73	109.75
10	shallow	Kherlen 7	48.15	108.63
10	deep	Kherlen 7	48.15	108.63
11	center	Kherlen 8	47.93	108.44
11	side	Kherlen 8	47.93	108.44
12	shallow	Little Missouri	46.33	-103.92
12	deep	Little Missouri	46.33	-103.92
13	shallow	Little Missouri 2	46.96	-103.51
13	deep	Little Missouri 2	46.96	-103.51
14	center	Little Missouri 3	45.43	-104.05
14	side	Little Missouri 3	45.43	-104.05
15	shallow	Little Missouri 4	45.55	-103.97
15	deep	Little Missouri 4	45.55	-103.97
16	shallow	Little Missouri 5	45.77	-103.89
16	deep	Little Missouri 5	45.77	-103.89

17	center	Niobrara	42.67	-99.77
17	side	Niobrara	42.67	-99.77
18	center	Niobrara 2	42.54	-99.71
18	side	Niobrara 2	42.54	-99.71
19	center	Niobrara 3	42.79	-100.07
19	side	Niobrara 3	42.79	-100.07
20	shallow	Niobrara 4	42.55	-100.11
20	deep	Niobrara 4	42.55	-100.11
21	center	Tensleep	44.25	-107.22
21	side	Tensleep	44.25	-107.22
22	shallow	Tensleep	44.25	-107.22
22	deep	Tensleep	44.25	-107.22
23	center	Tensleep	44.25	-107.22
23	side	Tensleep	44.25	-107.22
24	shallow	Zakhvan	48.31	93.49
24	deep	Zakhvan	48.31	93.49
25	center	Zakhvan 2	47.1	97.63

25	side	Zakhvan 2	47.1	97.63
26	center	Zakhvan 3	46.59	97.26
26	side	Zakhvan 3	46.59	97.26
27	center	Zakhvan 3	46.59	97.26
27	side	Zakhvan 3	46.59	97.26
28	center	Kherlen 6	47.73	109.75
28	side	Kherlen 6	47.73	109.75
29	center	Little Missouri	46.33	-103.92
29	side	Little Missouri	46.33	-103.92
30	center	Little Missouri 4	45.55	-103.97
30	side	Little Missouri 4	45.55	-103.97
31	center	Little Missouri 4	45.55	-103.97
31	side	Little Missouri 4	45.55	-103.97
32	center	Niobrara 4	42.55	-100.11
32	side	Niobrara 4	42.55	-100.11
33	center	Delgermurun	49.64	99.92
33	side	Delgermurun	49.64	99.92

34	center	Zakhvan 4	48.28	93.48
34	side	Zakhvan 4	48.28	93.48
35	shallow	Zakhvan 3	46.59	97.26
35	deep	Zakhvan 3	46.59	97.26
36	center	Eg 2	50.1	101.59
36	side	Eg 2	50.1	101.59
37	center	Kherlen 3	48.05	113.63
37	side	Kherlen 3	48.05	113.63

Supplemental Table 2.3. Metabolism estimates and standard deviations with manipulated barometric pressure illustrated in Figure 2.10

Sensor	Pressure	GPP	GPP sd	ER	ER sd	K	K sd
	atm	mg O_2 L^{-1} day^{-1}				day^{-1}	
collected	0.732	0.933	0.140	4.493	0.389	9.487	0.673
variable	0.728	0.813	0.121	3.868	0.306	8.986	0.549
daily mean	0.728	0.993	0.181	4.313	0.475	9.877	0.874
monthly mean	0.729	0.963	0.156	4.315	0.413	9.712	0.749
sea level	0.999	0.017	0.016	8.674	0.219	2.835	0.070
station	0.861	0.151	0.063	7.276	0.369	4.293	0.187
plus 500m	0.689	1.565	0.279	0.851	0.233	14.118	1.551

Supplemental Table 2.4. Metabolism estimates and standard deviations from Figure 2.2

Adjustment in mg L^{-1} (%)	GPP	GPP sd	ER	ER sd	K	K sd
	mg O_2 L^{-1} day^{-1}				day^{-1}	
-0.4 (-5.3)	1.05	0.14	-4.79	0.48	11.05	1.04
-0.2 (-2.7)	1.13	0.16	-3.08	0.36	11.77	1.09
-0.1 (-1.3)	1.18	0.16	-2.14	0.29	15.91	2.23
0 (0)	1.23	0.17	-1.13	0.21	7.03	1.60
0.1 (1.3)	1.36	0.16	-0.14	0.11	10.69	0.99
0.2 (2.7)	2.51	0.34	-0.03	0.03	10.35	0.93
0.4 (5.3)	2.01	0.37	-0.02	0.02	9.72	0.82
Coefficient of Variation	0.37	0.46	1.13	0.81	0.24	0.40

Supplemental Table 2.5. Papers reviewed from Web of Science Search

Authors	Title	Journal	DOI
Nebgen, EL; Herrman, KS	Effects of shading on stream ecosystem metabolism and water temperature in an agriculturally influenced stream in central Wisconsin, USA	ECOLOGICAL ENGINEERING	10.1016/j.ecoleng.2018.10.023
Atnoee, AB; Riis, T; Baattrup-Pedersen, A	Comparison of metabolic rates among macrophyte and nonmacrophyte habitats in streams	FRESHWATER SCIENCE	10.1086/687842
Houser, JN; Bartsch, LA; Richardson, WB; Rogala, JT; Sullivan, JF	Ecosystem metabolism and nutrient dynamics in the main channel and backwaters of the Upper Mississippi River	FRESHWATER BIOLOGY	10.1111/fwb.12617
Hornbach, DJ; Beckel, R; Hustad, EN; McAdam, DP; Roen, IM; Wareham, AJ	The influence of riparian vegetation and season on stream metabolism of Valley Creek, Minnesota	JOURNAL OF FRESHWATER ECOLOGY	10.1080/02705060.2015.1063096
Qasem, K; Vitousek, S; O'Connor, B; Hoellein, T	The effect of floods on ecosystem metabolism in suburban streams	FRESHWATER SCIENCE	10.1086/703459

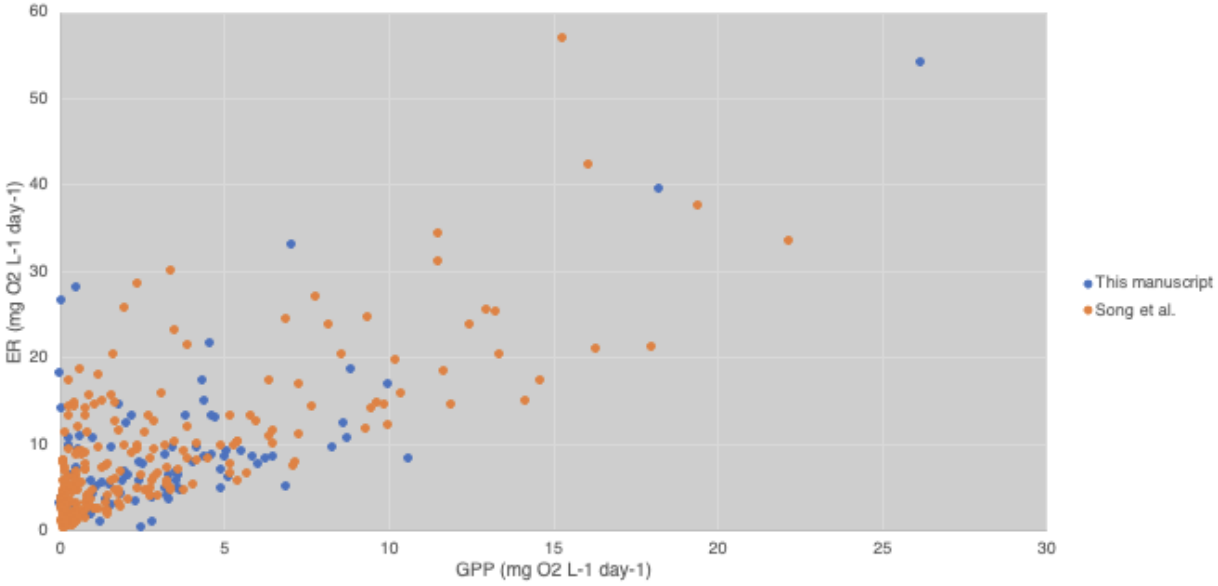
Koopmans, DJ; Berg, P	Stream oxygen flux and metabolism determined with the open water and aquatic eddy covariance techniques	LIMNOLOGY AND OCEANOGRAPHY	10.1002/lno.10103
Limburg, KE; Swaney, DP; Hall, MH	Temporal and spatial dynamics of a "Rust-Belt" urban stream: Metabolic and water quality responses to hardened land	URBAN ECOSYSTEMS	10.1007/s11252-012-0273-0
Schindler, DE; Jankowski, K; A'mar, ZT; Holtgrieve, GW	Two-stage metabolism inferred from diel oxygen dynamics in aquatic ecosystems	ECOSPHERE	10.1002/ecs2.1867
Staehr, PA; Asmala, E; Carstensen, J; Krause-Jensen, D; Reader, H	Ecosystem metabolism of benthic and pelagic zones of a shallow productive estuary: spatio-temporal variability	MARINE ECOLOGY PROGRESS SERIES	10.3354/meps12697
Mejia, FH; Fremier, AK; Benjamin, JR; Bellmore, JR; Grimm, AZ; Watson, GA; Newsom, M	Stream metabolism increases with drainage area and peaks asynchronously across a stream network	AQUATIC SCIENCES	10.1007/s00027-018-0606-z
Kupilas, B; Hering, D; Lorenz, AW; Knuth, C; Gucker, B	Hydromorphological restoration stimulates river ecosystem metabolism	BIOGEOSCIENCES	10.5194/bg-14-1989-2017
Tang, S; Sun, T; Shen, XM; Qi, M; Feng, ML	Modeling net ecosystem metabolism influenced by artificial hydrological regulation: An application to the Yellow River Estuary, China	ECOLOGICAL ENGINEERING	10.1016/j.ecoleng.2014.04.025
Hall, RO; Tank, JL; Baker, MA; Rosi-Marshall, EJ; Hotchkiss, ER	Metabolism, Gas Exchange, and Carbon Spiraling in Rivers	ECOSYSTEMS	10.1007/s10021-015-9918-1
Hood, JM; Benstead, JP; Cross, WF; Huryn, AD; Johnson, PW; Gislason, GM; Junker, JR; Nelson, D; Olafsson, JS; Tran, C	Increased resource use efficiency amplifies positive response of aquatic primary production to experimental warming	GLOBAL CHANGE BIOLOGY	10.1111/gcb.13912
Song, C; Dodds, WK; Trentman, MT; Ruegg, J; Ballantyne, F	Methods of approximation influence aquatic ecosystem metabolism estimates	LIMNOLOGY AND OCEANOGRAPHY-METHODS	10.1002/lom3.10112

Fuss, T; Behounek, B; Ulseth, AJ; Singer, GA	Land use controls stream ecosystem metabolism by shifting dissolved organic matter and nutrient regimes	FRESHWATER BIOLOGY	10.1111/fwb.12887
Venkiteswaran, JJ; Schiff, SL; Taylor, WD	Linking aquatic metabolism, gas exchange, and hypoxia to impacts along the 300-km Grand River, Canada	FRESHWATER SCIENCE	10.1086/683241
Sergeant, CJ; Bellmore, R; McConnell, C; Moore, JW	High salmon density and low discharge create periodic hypoxia in coastal rivers	ECOSPHERE	10.1002/ecs2.1846
Payn, RA; Hall, RO; Kennedy, TA; Poole, GC; Marshall, LA	A coupled metabolic-hydraulic model and calibration scheme for estimating whole-river metabolism during dynamic flow conditions	LIMNOLOGY AND OCEANOGRAPHY-METHODS	10.1002/lom3.10204
Aspray, KL; Holden, J; Ledger, ME; Mainstone, CP; Brown, LE	Organic sediment pulses impact rivers across multiple levels of ecological organization	ECOHYDROLOGY	10.1002/eco.1855
Guillon, S; Thorel, M; Flipo, N; Oursel, B; Claret, C; Fayolle, S; Bertrand, C; Rapple, B; Piegay, H; Olivier, JM; Vienney, A; Marmonier, P; Franquet, E	Functional classification of artificial alluvial ponds driven by connectivity with the river: Consequences for restoration	ECOLOGICAL ENGINEERING	10.1016/j.ecoleng.2018.12.018
Shen, XM; Sun, T; Su, MR; Dang, Z; Yang, ZF	Short-term response of aquatic ecosystem metabolism to turbidity disturbance in experimental estuarine wetlands	ECOLOGICAL ENGINEERING	10.1016/j.ecoleng.2019.06.005
Genzoli, L; Hall, RO	Shifts in Klamath River metabolism following a reservoir cyanobacterial bloom	FRESHWATER SCIENCE	10.1086/687752
Rodriguez-Castillo, T; Estevez, E; Gonzalez-Ferreras, AM; Barquin, J	Estimating Ecosystem Metabolism to Entire River Networks	ECOSYSTEMS	10.1007/s10021-018-0311-8
Siders, AC; Larson, DM; Ruegg, J; Dodds, WK	Probing whole-stream metabolism: influence of spatial heterogeneity on rate estimates	FRESHWATER BIOLOGY	10.1111/fwb.12896

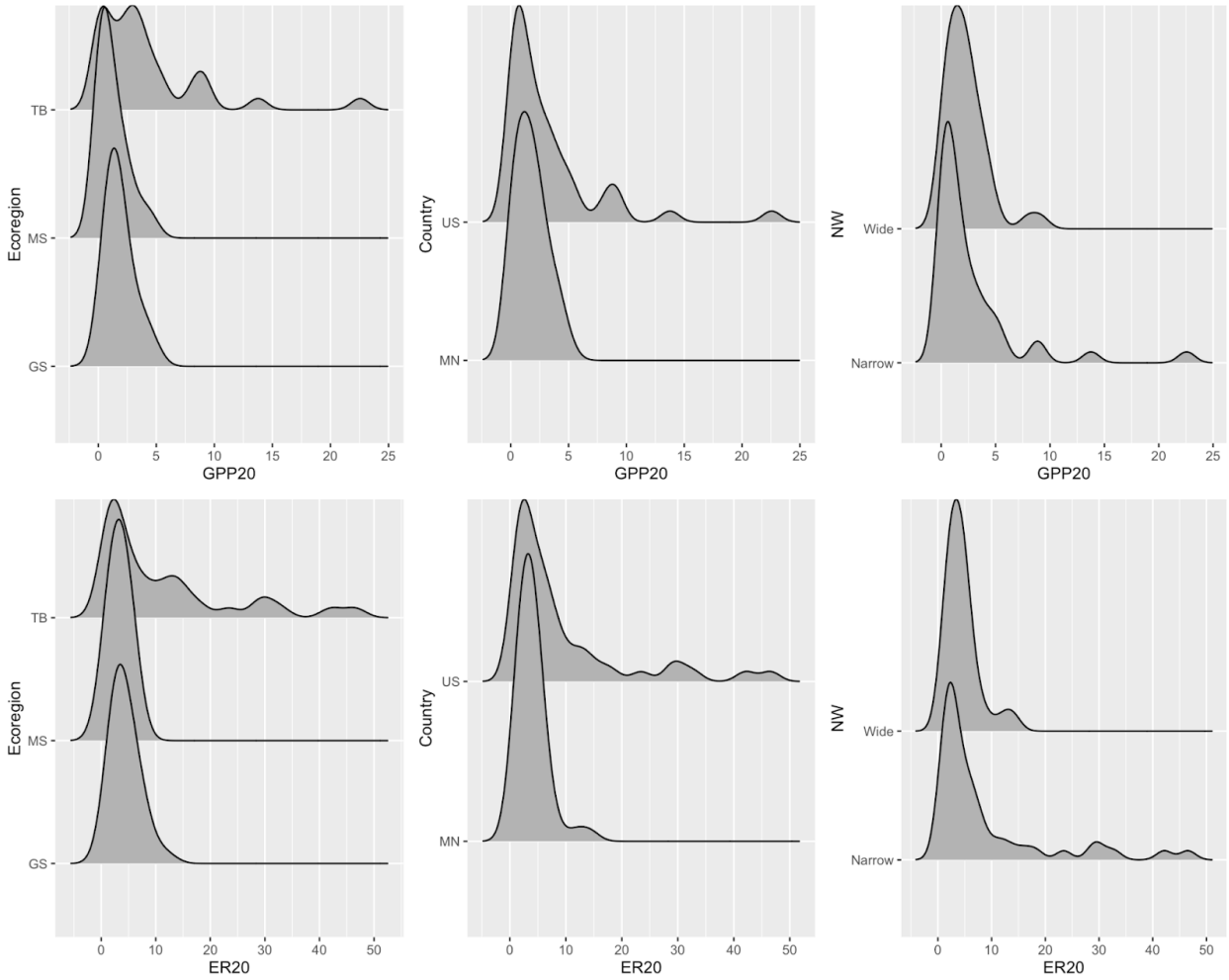
Reale, JK; Van Horn, DJ; Condon, KE; Dahm, CN	The effects of catastrophic wildfire on water quality along a river continuum	FRESHWATER SCIENCE	10.1086/684001
Blaszczak, JR; Delesantro, JM; Urban, DL; Doyle, MW; Bernhardt, ES	Scoured or suffocated: Urban stream ecosystems oscillate between hydrologic and dissolved oxygen extremes	LIMNOLOGY AND OCEANOGRAPHY	10.1002/lno.11081
Arroita, M; Elosegi, A; Hall, RO	Twenty years of daily metabolism show riverine recovery following sewage abatement	LIMNOLOGY AND OCEANOGRAPHY	10.1002/lno.11053
Hornbach, D; Hove, M; Agata, M; Albright, E; Cavazos, E; Friedman, C; Jay, K; Johnson, E; Johnson, K; Staudenmaier, A	Ecosystem structure and function in two branches of an eastern Minnesota, USA, trout stream	JOURNAL OF FRESHWATER ECOLOGY	10.1080/02705060.2016.1212118
Kallenbach, EMF; Sand-Jensen, K; Morsing, J; Martinsen, KT; Kragh, T; Raulund-Rasmussen, K; Baastrup-Spohr, L	Early ecosystem responses to watershed restoration along a headwater stream	ECOLOGICAL ENGINEERING	10.1016/j.ecoleng.2018.03.005
Wang, SR; Di Iorio, D; Cai, WJ; Hopkinson, CS	Inorganic carbon and oxygen dynamics in a marsh-dominated estuary	LIMNOLOGY AND OCEANOGRAPHY	10.1002/lno.10614
Attard, KM; Rodil, IF; Glud, RN; Berg, P; Norkko, J; Norkko, A	Seasonal ecosystem metabolism across shallow benthic habitats measured by aquatic eddy covariance	LIMNOLOGY AND OCEANOGRAPHY LETTERS	10.1002/lol2.10107
Graeber, D; Lorenz, S; Poulsen, JR; Heinz, M; von Schiller, D; Gucker, B; Gelbrecht, J; Kronvang, B	Assessing net-uptake of nitrate and natural dissolved organic matter fractions in a revitalized lowland stream reach	LIMNOLOGICA	10.1016/j.limno.2017.10.003
Hartwig, M; Borchardt, D	**Alteration of key hyporheic functions through biological and physical clogging along a nutrient and fine-sediment gradient**	**ECOHYDROLOGY**	**10.1002/eco.1571**
Hall, RO; Yackulic, CB; Kennedy, TA; Yard, MD;	Turbidity, light, temperature, and hydropeaking control primary	LIMNOLOGY AND OCEANOGRAPHY	10.1002/lno.10031

Rosi-Marshall, EJ; Voichick, N; Behn, KE	productivity in the Colorado River, Grand Canyon		
Pennington, R; Argerich, A; Haggerty, R	Measurement of gas-exchange rate in streams by the oxygen-carbon method	FRESHWATER SCIENCE	10.1086/698018
Alnoee, AB; Riis, T; Andersen, MR; Baattrup-Pedersen, A; Sand-Jensen, K	Whole-stream metabolism in nutrient-poor calcareous streams on Oland, Sweden	AQUATIC SCIENCES	10.1007/s00027-014-0380-5
Tromboni, F; Dodds, WK; Neres-Lima, V; Zandona, E; Moulton, TP	Heterogeneity and scaling of photosynthesis, respiration, and nitrogen uptake in three Atlantic Rainforest streams	ECOSPHERE	10.1002/ecs2.1959
Bortolotti, LE; St Louis, VL; Vinebrooke, RD; Wolfe, AP	**Net Ecosystem Production and Carbon Greenhouse Gas Fluxes in Three Prairie Wetlands**	ECOSYSTEMS	10.1007/s10021-015-9942-1
Saunders, WC; Bouwes, N; McHugh, P; Jordan, CE	A network model for primary production highlights linkages between salmonid populations and autochthonous resources	ECOSPHERE	10.1002/ecs2.2131
Riis, T; Levi, PS; Baattrup-Pedersen, A; Jeppesen, KG; Leth, SR	Experimental drought changes ecosystem structure and function in a macrophyte-rich stream	AQUATIC SCIENCES	10.1007/s00027-017-0536-1
Qin, QB; Shen, J	Pelagic contribution to gross primary production dynamics in shallow areas of York River, VA, USA	LIMNOLOGY AND OCEANOGRAPHY	10.1002/lno.11129
Lupon, A; Marti, E; Sabater, F; Bernal, S	**Green light: gross primary production influences seasonal stream Nexport by controlling fine-scale N dynamics**	ECOLOGY	10.1890/14-2296.1

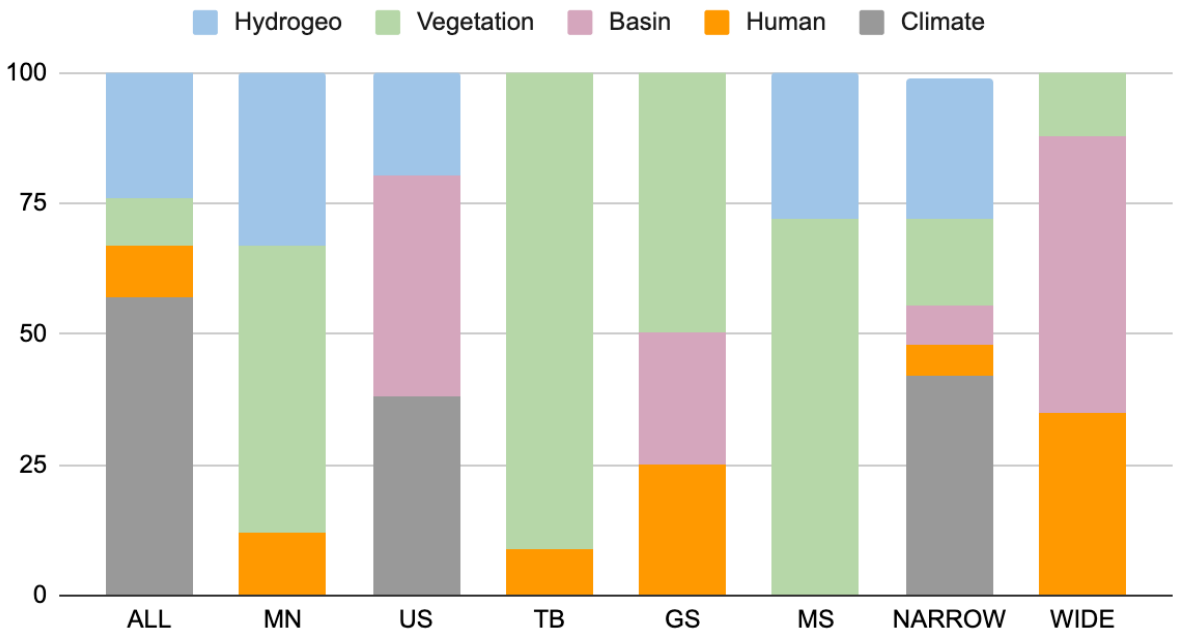
Appendix B - Supplemental Information for Chapter 3



Supplemental Figure 3.1. GPP by ER for our data (blue) and previously aggregated by Song et al. 2018.



Supplementary Figure 3.2. GPP and ER by Ecoregion, Country, and Narrow/Wide Designation.



Supplementary Figure 3.3. Summary representation of model primary explanatory variables

Appendix C - Supplemental Information for Chapter 4

Supplementary Table 4.1. Specific projection estimates of predicted rate and percent change by individual site within ecoregion

		Percent Change			Predicted Rate Estimate with Confidence Interval, in mg O_2 m^{-2} day^{-1}											
		Between temperature projections:			Fit	Upper	Lower	Fit	Upper	Lower	Fit	Upper	Lower	Fit	Upper	Lower
Subset	Site	0-1	0-1.5	0-2	0	0	0	1	1	1	2	2	2	2	2	2
US TB GPP	BEAUC1	-10.40	-14.41	-19.59	6.68	10.63	2.90	5.98	9.82	1.94	5.71	9.51	1.85	5.37	9.30	1.48
	BEAUC2	-103.97	-147.20	-184.68	0.54	5.45	-3.99	-0.02	4.74	-4.55	-0.26	4.45	-5.18	-0.46	3.70	-5.28
	BEAUC3	-189.45	-288.77	-371.07	5.09	10.14	0.61	-4.56	1.28	-10.26	-9.62	-3.17	-16.41	-13.81	-6.46	-21.30
	BEAUW1	-132.57	-202.73	-265.71	13.67	18.55	8.66	-4.45	4.04	-12.78	-14.05	-3.54	-24.67	-22.66	-9.23	-36.25
	BEAUW3	-10.08	-9.86	-15.46	7.61	11.82	3.70	6.84	11.36	3.12	6.86	11.24	2.49	6.43	10.39	1.95
	CARLC1	-17.68	-20.32	-31.62	4.25	8.24	0.15	3.50	7.45	-0.47	3.38	7.31	-0.85	2.90	6.90	-1.37
	CARLC2	-15.52	-18.36	-31.87	4.20	8.13	0.32	3.55	7.57	-0.57	3.43	6.97	-0.66	2.86	6.81	-1.09
	CARLC3	-22.45	-27.13	-31.70	4.13	8.02	0.33	3.21	7.63	-0.84	3.01	6.96	-0.89	2.82	6.69	-1.39
	CARLW2	-14.99	-26.44	-32.69	3.84	7.98	-0.18	3.26	7.55	-0.89	2.82	6.84	-1.26	2.58	6.73	-1.71
	CARLW3	-23.09	-30.28	-33.94	3.80	7.48	-0.42	2.92	7.31	-1.20	2.65	7.05	-1.46	2.51	6.69	-2.05
	CARUC3	-17.19	-33.89	-50.96	2.30	6.25	-1.63	1.90	5.62	-2.08	1.52	5.77	-2.52	1.13	5.14	-3.03
	CARUW2	-23.70	-25.72	-37.42	3.13	7.12	-0.95	2.39	6.39	-1.49	2.33	6.21	-1.37	1.96	5.71	-2.30
	CARUW3	-29.79	-50.75	-72.52	1.89	5.87	-2.06	1.33	5.28	-2.63	0.93	5.28	-3.10	0.52	4.79	-3.58
	HUMLC1	-13.74	-17.22	-25.73	5.38	9.28	1.42	4.64	8.59	0.81	4.46	8.41	0.45	4.00	7.95	0.16
	HUMLC2	-12.55	-16.77	-22.09	5.15	9.18	0.79	4.50	8.91	0.31	4.29	8.37	0.31	4.01	7.77	0.01

	HUMUC1	-10.89	-14.18	-17.28	7.27	11.22	3.16	6.48	10.64	2.49	6.24	10.05	2.06	6.01	10.01	1.76
	HUMUC2	-41.87	-62.44	-83.63	22.64	27.75	17.78	13.16	17.12	8.94	8.50	12.65	4.05	3.71	8.52	-1.19
	HUMUC3	-5.11	-11.60	-17.10	7.26	11.27	3.18	6.89	10.93	2.72	6.42	10.31	2.37	6.02	9.82	1.88
	Average	-17.44	-26.08	-32.28												
US TB ER	BEAUC1	9.76	14.65	21.50	17.33	29.11	5.96	19.02	31.43	6.91	19.87	31.42	7.97	21.05	33.66	9.55
	BEAUC2	6.45	7.33	17.02	15.14	27.21	2.79	16.11	28.31	3.98	16.25	28.28	4.05	17.71	30.20	6.11
	BEAUC3	8.89	14.81	15.14	19.93	32.61	7.69	21.71	34.59	9.29	22.88	34.99	9.65	22.95	35.73	9.99
	BEAUW1	5.42	11.20	12.21	20.07	32.39	7.88	21.16	34.56	7.63	22.32	34.42	8.34	22.53	36.38	9.87
	BEAUW3	8.53	16.29	16.80	19.93	33.27	7.53	21.63	34.06	9.83	23.17	34.23	11.27	23.27	35.18	11.60
	CARLC1	25.12	22.07	41.12	7.06	21.25	-6.15	8.84	22.44	-5.31	8.62	23.13	-4.66	9.97	23.66	-2.67
	CARLC2	31.76	48.40	79.26	5.04	17.57	-6.30	6.64	19.67	-5.17	7.48	20.01	-5.86	9.03	21.72	-3.89
	CARLC3	11.43	15.51	14.90	18.14	31.01	6.36	20.22	33.32	6.82	20.95	35.01	7.47	20.84	34.59	8.07
	CARLW2	28.22	31.18	40.93	7.67	19.65	-5.59	9.84	22.72	-2.89	10.06	21.77	-2.58	10.81	22.95	-1.88
	CARLW3	23.83	22.51	40.90	8.14	20.48	-4.17	10.07	22.74	-2.76	9.97	22.35	-2.41	11.46	25.28	-2.49
	CARUC3	3.42	7.62	12.16	29.13	41.13	17.47	30.13	43.21	17.96	31.35	44.22	18.53	32.68	46.05	19.41
	CARUW2	12.43	34.31	55.44	5.91	18.02	-6.64	6.64	19.22	-6.43	7.93	20.88	-5.09	9.18	21.40	-4.68
	CARUW3	3.68	6.72	9.14	35.36	48.72	22.60	36.67	49.80	22.57	37.74	50.53	23.43	38.60	52.97	24.37
	HUMLC1	51.04	73.54	99.04	3.34	16.32	-8.96	5.05	17.16	-8.06	5.80	18.60	-6.97	6.65	19.32	-5.37
	HUMLC2	84.05	128.65	116.99	3.10	15.92	-9.33	5.71	20.12	-8.52	7.09	20.98	-7.52	6.73	21.83	-7.80
	HUMUC1	14.01	19.08	29.11	10.99	23.42	-1.72	12.53	24.42	0.18	13.09	24.66	1.25	14.19	27.02	1.48
	HUMUC2	3.61	5.48	5.16	48.48	61.94	35.12	50.23	62.65	35.69	51.14	64.55	36.96	50.98	65.47	36.51
	HUMUC3	3.41	8.36	9.59	29.54	41.39	17.17	30.55	42.96	18.60	32.01	44.23	19.89	32.38	45.17	19.73
	Average	10.59	15.90	19.26												

MN TB GPP	KVDUC1	52.73	19.66	-42.47	-0.21	2.51	-2.85	-0.10	2.59	-2.60	-0.17	2.50	-2.89	-0.30	2.50	-2.80
	KVDUC3	26.92	9.45	-18.44	-0.28	2.48	-2.84	-0.21	2.61	-2.71	-0.26	2.40	-3.00	-0.33	2.24	-2.90
	KVDUW1	-12.26	-9.94	-65.00	0.23	2.92	-2.33	0.20	2.93	-2.25	0.21	2.84	-2.42	0.08	2.65	-2.44
	KVDUW2	73.42	-65.26	-85.65	-0.13	2.43	-2.84	-0.03	2.57	-2.66	-0.21	2.49	-2.92	-0.24	2.48	-2.96
	KVDUW4	81.52	94.05	-69.67	0.13	2.68	-2.60	0.24	2.86	-2.22	0.26	2.86	-2.49	0.04	2.68	-2.44
	ZAKLW1	-59.98	-82.80	-108.89	1.54	4.05	-1.00	0.62	3.16	-1.66	0.27	2.60	-2.19	-0.14	2.24	-2.59
	ZAKUC1	-2.42	-6.37	-10.40	2.13	4.61	-0.65	2.08	4.77	-0.58	2.00	4.42	-0.65	1.91	4.48	-0.80
	ZAKUC2	-7.99	-14.89	-14.87	2.94	5.57	0.11	2.70	5.38	0.02	2.50	5.22	-0.16	2.50	5.08	-0.30
	ZAKUC3	-9.17	-14.21	-17.89	3.26	6.07	0.41	2.96	5.71	0.06	2.80	5.52	-0.05	2.68	5.28	-0.13
	ZAKUC4	-7.77	-12.02	-13.66	1.68	4.11	-0.91	1.55	4.29	-1.01	1.48	4.01	-1.08	1.45	3.97	-1.14
	ZAKUC5	-11.44	-18.60	-24.02	2.17	4.55	-0.52	1.92	4.60	-0.66	1.77	4.19	-0.81	1.65	4.14	-0.84
	ZAKUW1	-3.80	-8.01	-7.45	2.47	5.14	-0.24	2.38	5.18	-0.29	2.28	4.90	-0.47	2.29	4.90	-0.38
	ZAKUW2	-7.21	-7.94	-10.33	2.55	5.08	-0.25	2.37	5.14	-0.21	2.35	4.97	-0.41	2.29	4.90	-0.35
	ZAKUW4	-4.19	-14.05	-14.12	1.86	4.29	-0.71	1.78	4.30	-0.81	1.59	3.96	-1.08	1.59	4.11	-0.99
	Average	-5.70	-10.98	-18.17												
MN TB ER	KVDUC1	-84.63	-128.73	-167.98	0.95	3.57	-2.12	0.15	3.01	-3.02	-0.27	3.16	-3.60	-0.65	2.73	-4.24
	KVDUC3	11.49	14.67	19.26	11.02	14.48	7.73	12.29	15.89	8.23	12.64	16.36	8.50	13.15	17.33	8.61
	KVDUW1	-214.04	-349.31	-418.13	0.65	3.70	-2.47	-0.74	2.13	-3.99	-1.62	2.04	-4.88	-2.06	1.59	-5.88
	KVDUW2	-108.32	-203.61	-266.66	0.59	3.44	-2.32	-0.05	3.14	-3.51	-0.61	2.95	-3.72	-0.98	2.88	-4.62
	KVDUW4	0.56	-14.26	-24.82	1.58	4.37	-1.23	1.58	4.37	-1.29	1.35	4.34	-1.83	1.18	4.41	-1.87
	ZAKLW1	-29.34	-32.97	-46.39	2.14	4.96	-0.93	1.51	4.74	-1.53	1.43	5.19	-1.97	1.14	4.90	-2.41
	ZAKUC1	-0.87	-2.44	2.57	3.00	5.99	0.07	2.97	5.61	0.16	2.92	5.80	-0.04	3.07	5.83	0.18
	ZAKUC2	9.87	14.57	19.68	14.18	17.32	10.81	15.58	18.87	12.15	16.24	19.62	12.96	16.97	20.54	13.24

	ZAKUC3	4.51	4.13	6.68	6.00	8.91	3.19	6.27	9.14	3.61	6.25	9.31	3.47	6.40	9.39	3.36
	ZAKUC4	37.57	49.12	66.66	1.25	4.40	-2.01	1.72	4.67	-1.37	1.87	4.98	-1.36	2.09	5.01	-0.97
	ZAKUC5	6.38	9.00	10.94	5.15	7.83	2.26	5.48	8.17	2.70	5.62	8.59	2.78	5.72	8.58	2.91
	ZAKUW1	-3.58	-4.35	-3.61	4.15	6.91	1.35	4.00	6.79	1.32	3.97	6.89	1.36	4.00	6.86	1.13
	ZAKUW2	-3.22	-2.95	-10.67	3.56	6.29	0.79	3.44	6.14	0.73	3.45	5.99	0.75	3.18	6.30	0.58
	ZAKUW4	2.94	2.20	0.90	4.50	7.32	1.81	4.63	7.19	1.85	4.60	7.48	1.79	4.54	7.48	1.74
	Average	-0.16	-2.70	-1.36												
US MS GPP	CLEUC2	14.95	34.77	25.69	0.28	1.36	-0.82	0.32	1.52	-0.86	0.38	1.60	-0.73	0.35	1.55	-0.81
	CLEUC3	26.20	124.74	163.41	0.09	1.23	-0.96	0.11	1.38	-1.03	0.20	1.44	-0.94	0.24	1.34	-0.92
	CRAUW1	8.17	2.59	14.41	0.95	2.15	-0.20	1.02	2.14	-0.14	0.97	2.14	-0.16	1.08	2.20	-0.13
	LAKUW1	24.15	15.41	36.20	0.43	1.70	-0.80	0.53	1.73	-0.78	0.50	1.89	-0.83	0.59	1.87	-0.77
	LBHUW1	0.85	2.78	2.20	3.99	5.25	2.85	4.03	5.37	2.79	4.10	5.29	2.96	4.08	5.46	2.80
	LBHUW2	1.90	3.15	2.19	2.96	4.09	1.76	3.01	4.16	1.93	3.05	4.11	1.89	3.02	4.23	1.91
	LBHUW3	2.41	4.33	2.45	2.14	3.33	0.94	2.20	3.39	0.95	2.24	3.52	1.00	2.20	3.57	1.03
	SOUUW1	274.58	339.13	490.95	-0.03	1.00	-1.11	0.06	1.16	-0.98	0.08	1.24	-1.04	0.13	1.24	-0.92
	TENUW1	9.31	11.79	8.89	0.90	2.06	-0.12	0.98	2.08	-0.17	1.00	2.06	-0.19	0.98	1.99	-0.02
	TENUW2	125.79	207.73	281.71	-0.06	1.13	-1.21	0.01	1.13	-1.13	0.06	1.20	-1.04	0.10	1.22	-1.07
	TONUC1	4.87	10.16	21.08	0.26	1.44	-0.91	0.27	1.55	-0.91	0.29	1.51	-0.94	0.32	1.56	-0.92
	TONUC2	2.09	14.20	12.29	0.73	1.78	-0.38	0.74	1.79	-0.35	0.83	1.87	-0.34	0.82	1.92	-0.41
	TONUC3	1.28	2.38	28.87	0.40	1.46	-0.76	0.41	1.57	-0.72	0.41	1.53	-0.67	0.52	1.57	-0.59
	Average	8.17	11.79	21.08												
US MS ER	CLEUC2	-2.49	-4.39	-3.87	1.61	3.29	-0.08	1.57	3.30	0.05	1.54	3.36	-0.25	1.54	3.37	-0.03
	CLEUC3	-16.92	9.64	2.16	0.38	2.27	-1.54	0.31	2.06	-1.53	0.41	2.14	-1.48	0.38	2.09	-1.38

	CRAUW1	-0.72	1.17	-2.05	5.57	7.60	3.64	5.53	7.34	3.74	5.63	7.43	3.77	5.45	7.34	3.64
	LAKUW1	-0.87	0.61	0.13	3.05	4.92	1.35	3.02	4.92	1.10	3.07	5.22	0.97	3.06	4.93	0.89
	LBHUW1	-3.47	-4.29	-3.78	3.12	4.74	1.29	3.01	4.75	1.24	2.99	4.74	1.30	3.00	4.63	1.31
	LBHUW2	2.04	1.34	-1.27	2.96	4.75	1.16	3.02	4.62	1.07	3.00	4.68	1.19	2.92	4.65	1.30
	LBHUW3	-4.06	-2.56	-1.01	2.15	4.05	0.21	2.07	4.00	0.24	2.10	4.18	0.16	2.13	4.10	0.06
	SOUUW1	-1.90	-3.49	-2.53	4.51	6.35	2.73	4.43	6.09	2.64	4.36	6.08	2.63	4.40	6.13	2.65
	TENUW1	-0.76	-1.95	-4.50	4.58	6.34	2.92	4.55	6.27	2.73	4.49	6.29	2.73	4.38	6.12	2.79
	TENUW2	-0.62	-1.77	-3.85	5.39	7.12	3.54	5.35	7.25	3.49	5.29	7.05	3.50	5.18	6.95	3.60
	TONUC1	-4.79	-1.89	-4.59	1.64	3.33	-0.23	1.56	3.29	-0.23	1.61	3.47	-0.21	1.56	3.28	-0.24
	TONUC2	3.14	3.04	-5.38	2.45	4.36	0.62	2.53	4.16	0.69	2.52	4.19	0.64	2.32	4.11	0.71
	TONUC3	-8.25	-3.52	-8.87	1.39	3.08	-0.38	1.27	3.00	-0.51	1.34	2.93	-0.41	1.27	2.97	-0.48
	Average	-1.90	-1.89	-3.78												
MN MS GPP	DELLW1	-3.36	-2.12	-3.94	2.94	4.16	1.73	2.84	3.99	1.57	2.88	4.22	1.61	2.83	4.09	1.53
	DELLW2	-2.02	-3.32	-5.66	3.98	5.19	2.71	3.90	5.01	2.71	3.84	5.00	2.57	3.75	5.04	2.56
	DELLW3	-11.74	-13.22	-14.63	1.33	2.34	0.28	1.17	2.26	0.12	1.16	2.22	0.14	1.14	2.28	0.07
	DELUC1	-12.39	-15.98	-11.80	0.56	1.61	-0.45	0.49	1.48	-0.63	0.47	1.54	-0.59	0.49	1.59	-0.68
	DELUC2	-8.02	-12.19	-22.61	0.59	1.72	-0.40	0.55	1.58	-0.57	0.52	1.62	-0.54	0.46	1.51	-0.62
	DELUC3	0.47	-1.04	-4.89	1.08	2.12	0.01	1.08	2.15	-0.02	1.07	2.04	0.02	1.03	2.07	-0.02
	DELUC4	-7.77	-8.65	-17.46	0.93	2.02	-0.11	0.86	1.87	-0.17	0.85	1.87	-0.19	0.77	1.78	-0.29
	EGILS1	-3.93	-5.26	-8.95	1.41	2.42	0.35	1.35	2.55	0.36	1.33	2.47	0.20	1.28	2.56	0.02
	EGILS2	-2.89	-7.92	-10.46	1.68	2.76	0.64	1.63	2.80	0.45	1.55	2.81	0.38	1.51	2.80	0.23
	EGILS3	-6.88	-21.89	-18.03	0.45	1.51	-0.69	0.41	1.69	-0.83	0.35	1.68	-0.95	0.36	1.80	-1.17
	EGILS4	-12.22	-8.35	-38.03	0.55	1.59	-0.53	0.48	1.68	-0.76	0.50	1.75	-0.77	0.34	1.69	-1.04

	EGILW1	-9.09	-11.71	-15.23	1.37	2.39	0.31	1.25	2.43	0.13	1.21	2.39	0.09	1.17	2.33	0.01
	Average	-7.32	-8.50	-13.21												
MN MS ER	DELLW1	14.05	20.67	28.97	5.61	7.52	3.65	6.39	8.53	4.25	6.77	9.06	4.63	7.23	9.43	4.94
	DELLW2	19.82	33.73	43.41	3.62	5.16	1.95	4.33	5.94	2.80	4.83	6.54	3.15	5.18	7.01	3.49
	DELLW3	23.51	35.19	45.38	3.71	5.35	2.08	4.58	6.23	2.89	5.02	6.77	3.10	5.40	7.24	3.52
	DELUC1	-84.28	-118.90	-175.78	1.80	3.62	0.06	0.28	2.66	-1.92	-0.34	2.46	-3.34	-1.36	2.16	-4.99
	DELUC2	-123.11	-166.30	-244.04	1.76	3.66	0.02	-0.41	2.07	-2.71	-1.17	1.80	-4.23	-2.54	1.69	-6.49
	DELUC3	33.48	52.68	68.29	3.85	5.56	2.22	5.14	6.87	3.44	5.88	7.71	3.94	6.48	8.65	4.58
	DELUC4	31.82	51.14	66.99	3.66	5.43	1.96	4.83	6.58	3.14	5.53	7.36	3.64	6.11	8.19	4.10
	EGILS1	10.73	19.77	22.38	3.76	5.34	2.15	4.16	6.01	2.37	4.50	6.21	2.59	4.60	6.65	2.74
	EGILS2	21.15	29.97	38.98	4.94	6.81	3.16	5.99	8.11	4.01	6.43	8.57	4.16	6.87	9.31	4.52
	EGILS3	0.15	9.86	3.95	3.10	4.93	1.39	3.11	5.25	1.00	3.41	5.70	1.04	3.23	5.82	0.80
	EGILS4	16.00	25.20	30.93	5.05	6.83	3.17	5.86	7.82	3.74	6.32	8.50	4.01	6.61	9.09	4.37
	EGILW1	-10.66	-10.09	-15.40	2.53	4.24	0.84	2.26	4.35	0.53	2.28	4.47	0.16	2.14	4.60	-0.43
	Average	15.02	22.94	29.95												
US GS GPP	LMRLC1	109.63	153.46	221.79	1.08	2.97	-0.78	2.26	3.87	0.60	2.74	4.44	1.04	3.48	5.32	1.44
	LMRLC2	71.57	114.65	165.08	1.42	3.22	-0.51	2.44	4.21	0.78	3.05	4.81	1.35	3.77	5.72	1.82
	LMRLC3	93.25	148.33	205.90	1.12	2.81	-0.85	2.16	3.88	0.45	2.78	4.46	1.03	3.42	5.33	1.55
	LMRUW1	57.02	90.99	125.13	1.87	3.62	0.00	2.93	4.64	1.30	3.56	5.33	1.69	4.20	6.18	2.06
	LMRUW2	56.36	91.10	129.40	1.86	3.68	-0.04	2.91	4.73	1.26	3.56	5.33	1.77	4.27	6.15	2.29
	LMRUW3	60.89	97.08	133.11	1.76	3.49	-0.17	2.84	4.61	1.23	3.48	5.14	1.64	4.11	6.14	2.07
	NIOLC1	107.85	159.26	223.26	1.08	2.85	-0.62	2.25	4.46	0.25	2.80	5.15	0.34	3.50	6.41	0.56
	NIOLC2	94.77	139.48	200.53	1.22	3.24	-0.76	2.37	5.01	-0.25	2.92	5.81	-0.02	3.66	6.80	0.28

	NIOLC3	48.71	74.04	100.74	2.44	4.45	0.50	3.62	6.06	1.19	4.24	6.91	1.11	4.89	8.06	1.43
	NIOLC5	129.51	184.15	265.09	0.92	2.58	-0.77	2.12	4.30	-0.09	2.63	5.11	0.13	3.38	6.14	0.39
	NIOLC6	65.33	95.18	134.46	1.86	3.64	0.07	3.08	5.38	0.85	3.63	6.21	0.81	4.36	7.25	1.41
	NIOLC7	47.80	69.61	105.98	2.31	4.18	0.35	3.42	5.84	1.07	3.93	6.64	1.28	4.77	7.85	1.45
	NIOLC8	179.46	276.73	403.40	0.64	2.71	-1.23	1.79	4.35	-0.61	2.42	5.18	-0.54	3.23	6.24	-0.10
	Average	71.57	114.65	165.08												
US GS ER	LMRLC1	75.07	103.49	142.72	1.32	4.05	-1.56	2.31	5.26	-0.55	2.69	5.81	-0.37	3.20	6.69	0.03
	LMRLC2	55.65	87.82	114.45	1.77	4.59	-0.95	2.75	5.62	0.20	3.32	5.99	0.47	3.79	6.94	0.56
	LMRLC3	92.24	132.43	159.63	1.13	4.12	-1.72	2.18	5.21	-0.76	2.63	6.05	-0.61	2.94	6.45	-0.42
	LMRUW1	11.06	16.26	21.25	9.65	12.96	6.42	10.71	13.87	7.49	11.22	14.61	7.86	11.70	14.88	8.26
	LMRUW2	62.02	81.61	114.79	1.89	4.70	-1.21	3.06	5.66	0.33	3.43	6.22	0.65	4.06	7.20	0.83
	LMRUW3	18.09	24.76	34.97	5.78	8.71	3.12	6.83	9.56	3.88	7.21	10.13	4.23	7.80	10.70	4.69
	NIOLC1	26.26	41.72	54.58	3.44	6.11	0.75	4.34	7.65	1.21	4.87	8.49	1.45	5.31	8.99	1.58
	NIOLC2	14.77	23.19	29.63	7.25	10.18	4.33	8.32	11.48	5.15	8.93	12.16	6.00	9.40	12.80	5.90
	NIOLC3	13.39	16.90	26.28	8.40	11.30	5.25	9.52	13.06	6.24	9.82	13.79	6.09	10.61	14.98	6.38
	NIOLC5	22.44	34.92	43.94	4.67	7.62	1.91	5.72	9.01	2.47	6.31	9.76	2.96	6.73	10.70	2.99
	NIOLC6	14.32	24.25	28.16	7.08	9.98	4.18	8.10	11.10	4.96	8.80	11.94	5.43	9.08	12.65	5.78
	NIOLC7	17.73	22.99	28.66	7.18	10.04	4.27	8.46	11.35	5.33	8.84	12.02	5.67	9.24	12.99	5.73
	NIOLC8	20.31	31.84	35.58	5.49	8.24	2.72	6.60	9.42	3.42	7.23	10.94	3.41	7.44	11.29	3.82
	Average	20.31	31.84	35.58												
MN GS GPP	KHELWM1	13.05	19.58	26.11	2.92	3.82	2.01	3.30	4.43	2.17	3.49	5.06	1.92	3.68	5.75	1.61
	KHELWM2	12.33	18.50	24.66	2.95	3.75	2.14	3.31	4.64	1.97	3.49	5.30	1.68	3.67	5.99	1.35
	KHELWM3	17.41	26.11	34.82	0.99	3.40	-1.41	1.17	4.33	-2.00	1.25	4.83	-2.32	1.34	5.32	-2.65

	KHELWM4	1048.61	1572.91	2097.21	0.38	2.66	-1.90	4.34	9.00	-0.32	6.32	12.92	-0.27	8.30	16.92	-0.32
	KHELWS1	11.30	16.95	22.60	3.37	4.65	2.09	3.75	6.02	1.49	3.94	6.74	1.14	4.13	7.47	0.79
	KHELWS2	10.87	16.31	21.74	3.26	4.58	1.94	3.62	5.91	1.32	3.79	6.61	0.97	3.97	7.32	0.62
	KHELWS3	4.28	6.42	8.56	1.73	3.09	0.36	1.80	3.88	-0.28	1.84	4.31	-0.64	1.87	4.77	-1.02
	KHELWS4	57.98	86.97	115.96	3.52	5.66	1.37	5.55	9.57	1.53	6.57	11.57	1.57	7.59	13.59	1.60
	KHEUCG1	35.02	52.53	70.04	0.99	2.96	-0.98	1.33	3.08	-0.41	1.51	3.39	-0.38	1.68	3.84	-0.48
	KHEUCG2	19.55	29.33	39.11	0.50	2.28	-1.29	0.59	2.07	-0.88	0.64	2.09	-0.81	0.69	2.21	-0.83
	KHEUCM1	12.63	18.94	25.25	3.02	3.85	2.18	3.40	4.74	2.05	3.59	5.41	1.77	3.78	6.12	1.44
	KHEUCM2	11.79	17.69	23.58	3.23	4.23	2.23	3.61	5.49	1.73	3.80	6.20	1.40	3.99	6.93	1.05
	KHEUCM3	248.21	372.31	496.41	0.72	3.01	-1.58	2.49	4.58	0.40	3.38	5.91	0.85	4.27	7.45	1.09
	KHEUWM1	133.53	200.30	267.06	0.93	2.16	-0.30	2.17	3.10	1.24	2.79	4.08	1.49	3.41	5.21	1.61
	KHEUWM2	13.37	20.05	26.74	2.85	3.86	1.84	3.23	4.23	2.23	3.42	4.82	2.02	3.61	5.49	1.73
	KHEUWM3	14.86	22.29	29.71	2.56	4.20	0.92	2.94	3.82	2.07	3.13	4.01	2.26	3.32	4.51	2.14
	KHEUWM4	8.05	12.08	16.10	1.90	2.97	0.82	2.05	2.94	1.16	2.12	3.25	1.00	2.20	3.67	0.73
	Average	13.37	20.05	26.74												
MN GS ER	KHELWM1	-11.79	-17.68	-23.57	5.13	6.60	3.66	4.52	5.87	3.17	4.22	5.68	2.77	3.92	5.57	2.27
	KHELWM2	-14.07	-21.11	-28.15	4.30	5.29	3.30	3.69	4.84	2.54	3.39	4.79	1.99	3.09	4.79	1.38
	KHELWM3	-19.30	-28.96	-38.61	3.13	5.18	1.09	2.53	5.26	-0.20	2.22	5.31	-0.86	1.92	5.38	-1.53
	KHELWM4	-17.78	-26.67	-35.56	3.40	4.21	2.59	2.80	4.15	1.44	2.49	4.19	0.80	2.19	4.25	0.13
	KHELWS1	-11.91	-17.87	-23.82	5.07	6.29	3.86	4.47	5.90	3.04	4.17	5.84	2.49	3.87	5.83	1.90
	KHELWS2	-11.96	-17.94	-23.92	5.06	6.27	3.84	4.45	5.91	2.99	4.15	5.86	2.43	3.85	5.86	1.84
	KHELWS3	-12.15	-18.22	-24.29	4.98	6.20	3.76	4.37	5.86	2.89	4.07	5.81	2.33	3.77	5.80	1.73
	KHELWS4	-11.10	-16.66	-22.21	5.44	6.68	4.21	4.84	6.27	3.40	4.54	6.21	2.87	4.24	6.19	2.28

	KHEUCG1	-18.09	-27.13	-36.18	3.34	4.18	2.50	2.74	3.56	1.91	2.44	3.50	1.37	2.13	3.51	0.76
	KHEUCG2	-18.05	-27.08	-36.11	3.35	4.28	2.41	2.74	3.55	1.93	2.44	3.45	1.44	2.14	3.43	0.85
	KHEUCM1	-37.35	-56.02	-74.70	1.62	3.06	0.18	1.01	2.94	-0.91	0.71	2.95	-1.53	0.41	2.98	-2.16
	KHEUCM2	-27.57	-41.36	-55.15	2.19	3.35	1.04	1.59	3.38	-0.21	1.29	3.44	-0.86	0.98	3.50	-1.53
	KHEUCM3	-27.31	-40.97	-54.63	2.21	3.50	0.92	1.61	3.06	0.16	1.31	2.97	-0.36	1.00	2.94	-0.94
	KHEUWM1	-16.02	-24.03	-32.03	3.77	4.52	3.03	3.17	4.07	2.27	2.87	4.06	1.68	2.57	4.09	1.04
	KHEUWM2	-19.25	-28.88	-38.50	3.14	3.87	2.41	2.54	3.54	1.53	2.23	3.54	0.93	1.93	3.58	0.28
	KHEUWM3	-14.81	-22.21	-29.62	4.08	5.16	3.01	3.48	4.16	2.80	3.18	3.94	2.41	2.87	3.88	1.87
	KHEUWM4	-15.25	-22.87	-30.50	3.96	4.83	3.10	3.36	4.12	2.60	3.06	4.04	2.08	2.76	4.04	1.47
	Average	-16.02	-24.03	-32.03												